AD-A096 691 DECISIONS AND DESIGNS INC MCLEAN VA F/G 15/4

INTELLIGENCE PREPARATION OF THE BATTLEFIELD: CRITIQUE AND RECOM--ETC(U)

FEB 81 L ADELMAN, M L DONNELL, R H PHELPS MDA903-79-C-0275

UNCLASSIFIED PR-81-4-304 NL

1 OF 1

AD A
096691

LEVEL

February 1981

FINAL REPORT PR 81-4-304

Intelligence Preparation of the Battlefield: Critique and Recommendations

Leonard Adelman
Michael L. Donnell
Decisions and Designs, Inc.

Ruth H. Phelps
U.S. Army Research Institute

THE VIEWS, OPINIONS, AND FINDINGS CONTAINED IN THIS REPORT ARE THOSE OF THE AUTHORS AND SHOULD NOT BE CONSTRUED AS AN OFFICIAL DEPARTMENT OF THE ARMY POSITION, POLICY, OR DECISION, UNLESS SO DESIGNATED BY OTHER OFFICIAL DOCUMENTATION.

DECISIONS and DESIGNS, INC.

0 53

Prepared for
U.S. Army Research Institute
Contract MDA903-79-C-0275

12

FINAL REPORT PR 81-4-304

INTELLIGENCE PREPARATION OF THE BATTLEFIELD: CRITIQUE AND RECOMMENDATIONS

by

Leonard Adelman and Michael L. Donnell
Decisions and Designs, Inc.

and

Ruth H. Phelps
U.S. Army Research Institute

Prepared for

U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, Virginia 22333
Contract MDA903-79-C-0275

February 1981

DECISIONS AND DESIGNS, INC.
Suite 600, 8400 Westpark Drive
P.O.Box 907
McLean, Virginia 22101
(703) 821-2828

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: 14 PR-81-4-304 ✓ — 2. GOVT ACCESSION NO.: AD-A096	3. RECIPIENT'S CATALOG NUMBER: 691
4. TITLE (and Subtitle): 6 INTELLIGENCE PREPARATION OF THE BATTLEFIELD: CRITIQUE AND RECOMMENDATIONS.	5. TYPE OF REPORT & PERIOD COVERED: 9 Final Report.
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): 10 Leonard/Adelman, Michael L./Donnell, Ruth H./Phelps 15	8. CONTRACT OR GRANT NUMBER(s): MDA903-79-C-0275
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Decisions and Designs, Inc., Suite 600, 8400 Westpark Drive, P. O. Box 907, McLean, VA 22101	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS: U. S. Army Research Institute, 5001 Eisenhower Avenue, Alexandria, VA 22333 11	12. REPORT DATE: February 1981
	13. NUMBER OF PAGES: 95
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): 12 95	15. SECURITY CLASS. (of this report): UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

For abridged version of this report, see PR 81-5-304

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Tactical intelligence	Decision analysis	Circular TC30-27
Combat intelligence	Decision aids	Circular TC34-3
Threat evaluation	Computer-aided analysis	
Decision theory	Judgment (psychology)	
Battlefields		

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Intelligence Preparation of the Battlefield (IPB), ~~outlined in Circulars TC30-27 and TC34-3~~ represents a major step toward improving procedures for correlating data with enemy intent. The purpose of this report is to show how IPB could be improved by incorporating decision-analytic techniques and computerized decision aids to improve the judgmental process inherent within IPB.

The first section of this report describes the necessary dependence of IPB

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

390 264

on human judgment. This section identifies the different types of judgment inherent in each of the five steps in the IPB analysis process proposed, ~~in TC30-27 and TC34-3.~~ The general conclusion is that except for terrain analysis and weather analysis, these circulars fail to tell intelligence analysts how they are to make the judgments necessary to implement the proposed IPB analysis process.

The second section of this report describes the results of research studying how well people make the judgments inherent in IPB. The research strongly suggests that people's ability to make these judgments can be improved substantially by training them in decision-analytic techniques and by giving them access to computerized judgment aids. Two judgment aids incorporating decision-analytic techniques are described in detail, for they are designed to assist Staff Intelligence Officers (G2/S2s) in looking at the battlefield from the perspective of the enemy commander. Since both aids provide displays designed to help analysts convey the reasons for their judgments to tactical commanders, they provide an adjunct to the graphic templating process proposed in IPB. Scientific research will improve (1) the cognitive skills and, thus, judgmental accuracy of intelligence analysts developing IPB templates, and (2) the communication process for conveying the reasons for these judgments to tactical commanders.

CONTENTS

FIGURES

INTELLIGENCE PREPARATION OF THE BATTLEFIELD: CRITIQUE AND RECOMMENDATIONS

1.0 SUMMARY

The purpose of this paper is to show how the Intelligence Preparation of the Battlefield (IPB) analysis process outlined in Circulars TC30-27 and TC34-3 could be improved by incorporating decision-analytic techniques and computerized decisions aids to improve the judgmental process inherent within IPB. To accomplish this purpose, the paper is divided into two major sections.

Section 2.0 describes the necessary dependence of IPB on human judgment. In particular, it provides a conceptual overview of the causal inference problem inherent in intelligence analysis, and identifies the different types of judgment inherent in each of the five steps in the IPB analysis process proposed in TC30-27 and TC34-3. The general conclusion is that except for terrain analysis (step 3) and weather analysis (step 4), these circulars fail to tell intelligence analysts how they are to make the judgments necessary to implement the proposed IPB analysis process.

Section 3.0 describes the results of research studying how well people make the judgments inherent in IPB. The research strongly suggests that people's ability to make these judgments can be improved substantially by training them in decision-analytic techniques and by giving them access to computerized judgment aids. Two judgment aids incorporating decision-analytic techniques are described in detail, for they are designed to assist Staff Intelligence Officers (G2/S2s) in looking at the battlefield from the perspective of

the enemy commander. Since both aids provide displays designed to help analysts convey the reasons for their judgments to tactical commanders, they provide an adjunct to the graphic templating process in IPB.

Due to the length of this report, the reader with limited time is referred to Section 1.0, which provides a detailed overview and summary. The reader who is only interested in the judgment-analytic critique of IPB (i.e., Section 2.0) is referred to Adelman, Donnell, and Phelps (1980).

1.1 IPB's Dependence on Judgment

The first section below describes the limitations in Circulars TC30-27 and TC34-3 regarding the necessary judgments in the following three steps of the proposed IPB analysis process: mission and threat evaluation (step 1), tactical intelligence zone determination and evaluation (step 2), and threat integration (step 5). The second section describes the different types of judgments involved in revising IPB templates, an area only minimally addressed in Circulars TC30-27 and TC34-3.

1.1.1 Basic analysis steps - The purpose of mission and threat evaluation (step 1) is to review order of battle holdings so as to identify the gaps in collected intelligence data, and thereby, guide future collection efforts. Bowen, et al. (1975) point out the judgmental nature of order of battle holdings. The intelligence team must evaluate a large amount of incoming data, much of which may be unreliable, in order to decide whether it "knows" the order of battle holdings. In addition, they must integrate the order of battle holdings, complete or not, into a judgment of enemy intent. Although the proposed IPB process uses templates to represent the output of this judgmental process,

it provides no means for representing, pictorially or otherwise, the bases for the judgments underlying threat evaluation. Furthermore, TC30-27 and TC34-3 fail to provide an explicit representation of the probable relations between actual enemy intent and the order of battle holdings that repesent indicators of intent; consequently, it offers no explicit rules or heuristics describing how order of battle holdings are to be combined into judgments of enemy intent. Scientific research on causal inference strongly suggests that threat evaluation judgments could be improved by providing general information about probable cause-effect relations.

Tactical intelligence zone determination and evaluation (step 2) is supposed to identify how the opposing forces (OPFOR) threat evaluated in step 1 should look in the battlefield, in general. Circulars TC30-27 and TC34-3 accomplish this, however, only to a limited degree. There is, for example, no discussion of the general indicators for different OPFOR threats. Yet research by Johnson, Spooner, and Jaarsma (1977) suggests that this would be a valuable addition for tactical commanders. They found that a sample of forty-three captains in the Intelligence Officers Advanced Course knew only nineteen of the forty-nine separate indicators listed in Field Manual 30-5, Combat Intelligence, for four of the OPFOR threats that are to be evaluated in step 1. TI zone determination and evaluation could be improved by indicating (1) the relative accuracy of individual indicators for different OPFOR courses of action, and (2) how to combine individual indicators into a global assessment of threat.

The objective of the fifth and final step in the proposed IPB analysis process is to relate "how the enemy would like to fight" to a specific terrain and weather scenario as the basis for determining "how the enemy might have

to fight" (TC34-3, p. 5-1). According to both circulars, this is to be accomplished through the use of situation, event, and decision support templates. These templates are proposed as a means of helping commanders "visualize" enemy capabilities in a particular combat setting. They require intelligence analysts to make a number of different kinds of judgments. Both circulars fail, however, to tell intelligence analysts how to make these judgments. These different judgments are identified below for each template, in turn. In addition, limitations in the circulars are identified, and suggestions are offered for their improvement.

The situation template shows how enemy forces probably would look within the different mobility corridors under consideration. Underlying the development of this template is a complex judgmental process. We quote TC30-27 on this point.

> "While the enemy commander may not have unlimited options as to possible courses of action, he will probably have enough options to make the analyst's job of determining probable courses of action extremely difficult. Situation templates are derived based on the best military judgment of the analyst" (p. 4-4).

TC30-27 and TC34-3, however, provide only minimal information on how analysts are to exercise their "... best military judgment ..." when developing the situation templates. If the situation template is to reflect military judgment about different OPFOR courses of action, it is not enough to show analysts how to perform a terrain and weather analysis. In addition, the circulars should identify the factors that are to be considered when making this judgment, for the analysis team will have to incorporate judgments about other, more ambiguous factors, such as perceived U.S. force

strength and risk, that an OPFOR commander would certainly consider when selecting a course of action. Furthermore, the circulars should discuss the trade-offs that OPFOR commanders are likely to make when evaluating the utility of different courses of action. Rarely will it be true that one course of action is better than all others on every factor. Enemy commanders will be forced to differentially weight aspects of their doctrine with the characteristics of the situation immediately facing them. TC30-27, TC34-3, or supporting documentation should indicate what these trade-offs are likely to be under different terrain and weather constraints, if the situation template is to accurately represent military judgment under different circumstances.

"The event template is a time and logic sequence of enemy tactical indicators or events which are keyed to a series of situational templates" (TC30-27, pp. 1-6). Although TC30-27 and TC34-3 provide no example of an event template, it appears that the template must identify the different types of information necessary to confirm the adoption of a particular course of action. Such information is, in fact, required to complete the "events analysis matrix" (see TC34-3, p. 5-7). To complete this matrix, analysts must indicate those "event activities" they expect to see, as well as when and where they expect to see them, for each avenue of approach. In doing this, the analysts are essentially making a series of conditional probability judgments. That is, they must say something like, "If the enemy actually took this particular course of action, then these indicators and events have a higher probability of being observed than others." The word "probability" must be emphasized because there is not a perfect relationship between indicators, events, and actual enemy intent. The enemy will be expected to use deceptive measures. In fact, "An integral part of templating is the consideration of deception events associated with each course of action" (TC30-27, p. 4-11). As a

result, the intelligence team will be forced to make conditional probability judgments about what indicators and events they think are most indicative of the OPFOR adoption of different courses of action. Neither TC30-27 nor TC34-3, however, provide any information about the probabilistic relations between different indications and different OPFOR threats, a point made previously when discussing TI zone determination and evaluation.

The third, and final, effort in threat integration is development of a decision support template. This template "... is used to illustrate enemy probable courses of action as the basis for comparing friendly courses of action" (TC30-27, p. 4-13). Described in this way, the decision support template "... is essentially the INTELLIGENCE ESTIMATE in graphic format" (TC30-27, p. 1-6). It represents the analysts' most up-to-date estimate about the relative likelihood of the enemy's potential courses of action. Again, however, TC30-27 and TC34-3 fail to indicate *how* the judgments underlying this graphic representation are to be made, or in fact are made, by individual analysts. Psychological research suggests that this may result in decision support templates that (1) are not as accurate as they could be, and (2) do not facilitate communication between analysts and commanders as much as they could otherwise.

1.1.2 Template revision - Up to this point there has been no discussion of the dynamic nature of the proposed IPB templating process. The templating process is not supposed to stop with the first decision support template. The decision support template at the end of one cycle of the process represents the situation template at the beginning of the next cycle. The event templates and event matrices are used to convey, as quickly as possible, revised estimates about the likelihood of different OPFOR courses of action. These revised estimates of enemy intent are represented in the

decision support template. "Used properly, they [i.e., the templates) provide for continuing integrative analysis of OPFOR capabilities, vulnerabilities and courses of action" (TC30-27, p. 1-5). Figure 1-1 provides a schematic representation of this iterative process.

The dynamic nature of the proposed templating process is heavily dependent on a complex judgment process. First, intelligence analysts using the proposed IPB analysis process must identify the initial hypotheses regarding possible enemy courses of action, and if possible, make an initial estimate of the most probable course of action. As discussed above, this requires consideration of how the enemy generally makes judgments of intent, in addition to an evaluation of how physical terrain and weather factors favor different courses of action. Second, intelligence analysts using the proposed IPB process must indicate the events that are likely to be observed for each course of action. The information within all the event templates represents, in qualitative terms, likelihood ratios that indicate how much more likely certain events are to be observed for one course of action than another. Third, the analysts must use the many pieces of potentially fallible data reported in the event analysis matrices (and other sources) to revise their initial hypotheses about the enemy's most likely course of action. The data reported in the event analysis matrices also represent conditional probabilities, for certain events have a higher probability of being observed only when the enemy is actually taking a particular course of action. The events reported in the event analysis matrices are matched with those hypothesized in the event templates in order to develop a revised estimate of enemy intent, which is represented in the decision support template. The decision support template is now the situation template for the next iteration.

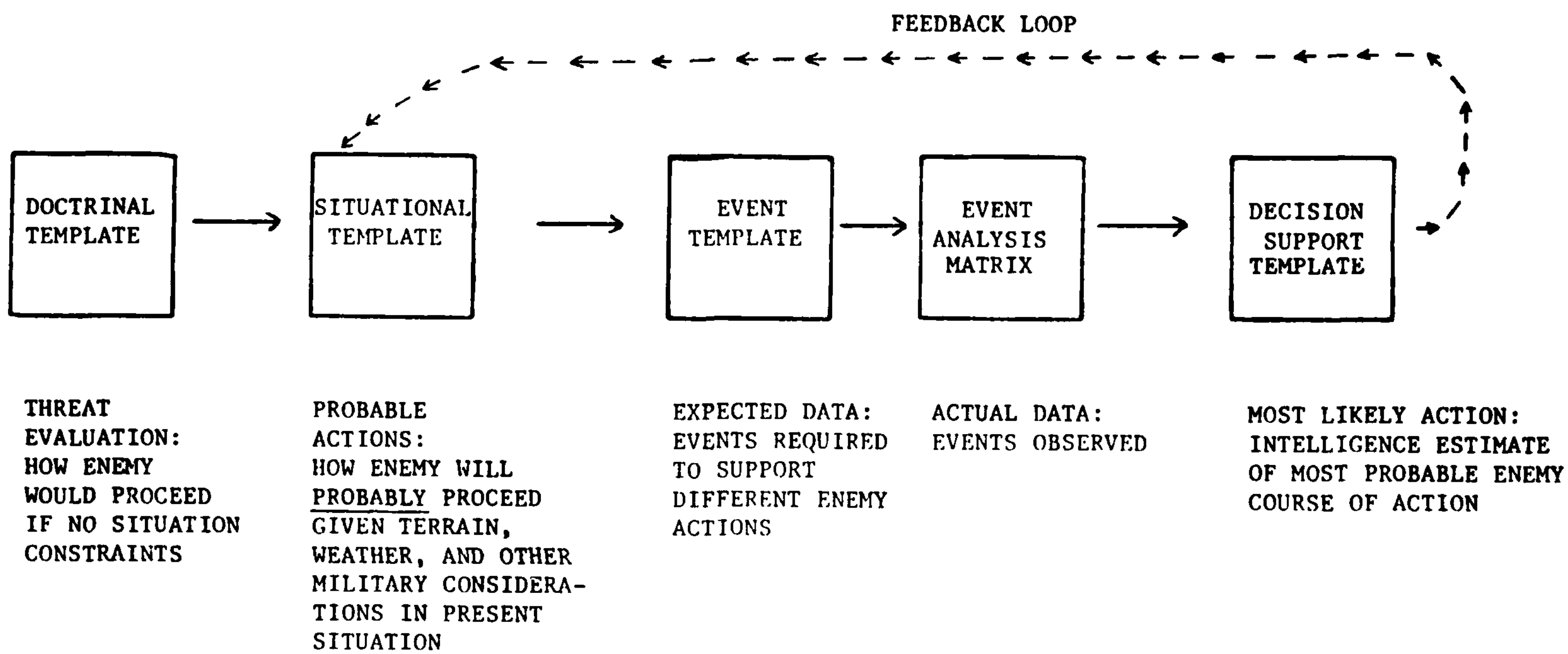

Figure 1-1

SCHEMATIC REPRESENTATION OF THE DYNAMIC NATURE OF THE PROPOSED IPB TEMPLATING PROCESS

1.2 Causal Inference Research Findings and Computer-based Judgment Aids Applicable to IPB

This section is divided into two parts. The first part describes causal inference research findings and computerized judgment aids relevant to the development of the initial decision support template. The second part describes research and aids relevant to the revision of decision support templates on the basis of newly collected data represented in the event template and event analysis matrix.

1.2.1 Research and aids applicable to the development of the initial decision support template - Research on causal inference with experts in different fields indicates that the level of judgmental accuracy and interpersonal agreement is a direct function of the characteristics of the judgmental task facing the experts. This conclusion is further supported by controlled laboratory research, which has shown that (1) judgmental accuracy and agreement can be increased or decreased by manipulating characteristics of the task, and (2) that relatively high levels of accuracy and agreement can be maintained under conditions that normally prevent them, by using judgment aids that provide persons with information about task characteristics.

Causal inference tasks can be characterized in terms of their formal characteristics. Formal task characteristics refer to the statistical relationships among task variables. Research on the following two task characteristics is of particular relevance to effective intelligence analysis: (a) the predictability of individual attributes, and (b) the total predictability of all the attributes in the inference task. Regarding the former, researchers have found achievement lowest in tasks in which individual attributes have equal predictability, which is frequently the case in reviewing Tactical Order of Battle according to

Bowen et al. (1975). Regarding the latter, researchers have found a positive linear function between achievement and overall task predictability; lower task predictability results in lower achievement. This occurs because people become less consistent in their judgments as the task becomes less predictable, even when they know how to perform. This finding is of particular importance to intelligence analysts because the generally low level of overall predictability in some intelligence tasks suggests that analysts may perform suboptimally because they do not have computerized aids for utilizing their knowledge of enemy activities in a rapid, highly consistent fashion.

Controlled laboratory research has shown that formal task characteristics also affect interpersonal agreement. Two task characteristics of particular relevance to intelligence analysis are overall task predictability and attribute intercorrelations. Regarding the former, the lower the level of task predictability, the lower the level of interpersonal understanding and agreement. This occurs because people try to predict the randomness (i.e., uncertainty) in the causal inference task; consequently, they make it difficult for the other person to learn how they make their judgments, and thereby, how to resolve the disagreement. Attribute intercorrelations also reduce interpersonal learning. For example, if two attributes (or factors) have a high positive correlation, people can rely on different attributes and still have high levels of accuracy and agreement. If the situation should change, however, such that the attributes are no longer correlated, the two persons will disagree. Furthermore, the dispute will be difficult to resolve because both participants will point to their recent successes, although only one of the individuals may actually know the valid indicator.

Laboratory research has shown that computerized judgment aids that quantitatively describe how people combine information on multiple attributes into an overall judgment greatly enhances judgmental skill and interpersonal agreement. Computer graphics devices have been used widely as judgment aids for presenting cognitive feedback, feedback that permits persons to compare the formal properties of their judgmental system with those of the task. Cognitive feedback provided by such displays has been shown to lead to both a faster rate of learning and a higher level of achievement (r_a) than outcome feedback in tasks that varied the (a) differences in the individual predictabilities of the attributes, and (b) the overall level of task predictability. While this finding is due in some part to greater knowledge acquisition with cognitive rather than outcome feedback, it is primarily due to higher cognitive control with cognitive feedback. This finding strongly supports our position that computerized judgment aids will improve the judgmental accuracy of experienced intelligence analysts using IPB by greatly increasing their ability to implement their knowledge. Furthermore, such aids will improve the training of analysts learning IPB.

Computerized, multi-attribute judgment aids enhance interpersonal learning because they help people overcome the inaccuracy and inconsistency of verbal self-reports. Verbal reports are often inaccurate because people inaccurately estimate the weight they place on various attributes. Research indicates that people tend to underestimate the weight they place on important factors and overestimate the weight they place on unimportant factors when compared to a quantitative analysis of their judgments. People often apply their judgments inconsistently because of task unpredictability (or uncertainty). Research has found that inconsistency leads to interpersonal misunderstanding and conflict, for the greater the inconsistency in one's judgment

the more difficult it is for others to learn how one makes one's judgment. Furthermore, research has shown that judgment aids that quantitatively describe (1) how different people weight the different attributes, (2) the overall level of consistency for each person's judgments, and (3) the implications of interpersonal differences in judgment, will enhance interpersonal understanding and reduce interpersonal conflict.

Multi-attribute utility assessment (MAUA) techniques and judgment aids can be used to help intelligence analysts develop the initial decision support template. They provide (1) a logically defensible conceptual structure for structuring the factors used in estimating enemy intent, (2) techniques for estimating the necessary trade-offs inherent in estimating enemy intent, (3) an analytical procedure for combining the multiple trade-off judgments into an overall assessment of the likelihood of different enemy courses of actions, (4) a means of systematically investigating the implications of differences of opinion in judgment, and (5) a means of conveying all this information pictorially, thus providing an important adjunct to the proposed pictorial format of the decision support template.

A computerized judgment aid called ENCOA (Enemy Courses of Action) has been developed to assist Staff Intelligence Officers (G2/S2s) develop the initial decision support template; see Patterson and Phelps (1980) for a complete description. ENCOA provides analysts with a systematic procedure for evaluating each potential OPFOR course of action on twenty-four factors affecting enemy intent. The factors incorporated into ENCOA not only include the order of battle, terrain, and weather factors discussed in circulars TC30-27 and TC34-3, but other factors that any tactical commander would consider when selecting a course of action. These

factors are grouped into the five categories indicated in Figure 1-2.

When developing the initial decision support template, intelligence analysts can use ENCOA to assess each OPFOR course of action on each of the factors comprising the five groups of factors. Thus, each course of action would receive a score on Terrain, U.S. Force Factors, OPFOR Force Factors, Weather Factors, and Risk Factors, where the score represented the utility of the course of action to an OPFOR commander. Then, these scores can be differentially weighted to represent the trade-offs of the OPFOR commander. The scores and weights can be combined to predict the overall utility of each course of action to the OPFOR commander. In addition, ENCOA can be used to quantitatively and pictorially present the scores, weights, overall utilities, and subsequent sensitivity analyses to the friendly commander. Such information describing (1) how the analysts reached their conclusions about the most likely OPFOR courses of action, as well as (2) the implications of differences in opinion between the analysts, represent an important adjunct to the decision support template.

1.2.2 <u>Research and aids applicable to the revision of decision support templates</u> - The iterative, templating process proposed in circulars TC30-27 and TC34-3 can be represented quantitatively by Bayes' Theorem, which is shown in equation [1].

$$\frac{P(H_1)}{P(H_2)} \times \frac{P(D|H_1)}{P(D|H_2)} = \frac{P(H_1|D)}{P(H_2|D)} \qquad [1]$$

(Prior Probabilities)	(Conditional Probabilities)	(Posterior Probabilities)
Situational Template	Event Templates; Event Analysis Matrices	Decision Templates

I. Terrain Factors

As related to mission accomplishment and considering current OPFOR doctrine, score each OPFOR course of action in terms of how well it:

1.1 Exploits field of fire afforded by terrain features.

1.2 Exploits cover and concealment afforded by terrain features.

1.3 Exploits mobility provisions due to terrain features.

1.4 Accomplishes rapid seizure or denial of key terrain.

1.5 Exploits observation provisions of terrain.

1.6 Exploits or accommodates natural and artificial obstacles.

II. U.S. Force Factors

As related to mission accomplishment and considering current U.S. doctrine, score each OPFOR course of action in terms of how well it exploits what you know or estimate about:

2.1 U.S. disposition.

2.2 U.S. strength and condition.

2.3 U.S. reserves.

2.4 U.S. logistic support.

2.5 Probable U.S. actions/reactions.

2.6 U.S. command and control capabilities/ vulnerabilities.

III. Opposing Force Factors

As related to mission accomplishment and considering current OPFOR doctrine, score each OPFOR course of action in terms of how well it exploits or accommodates:

3.1 OPFOR current disposition.

3.2 OPFOR strength and condition.

3.3 OPFOR reserves.

3.4 OPFOR logistic support.

3.5 OPFOR command and control capabilities/ vulnerabilities.

IV. Weather Factors

As related to mission accomplishment, score each OPFOR course of action in terms of how well it exploits:

4.1 Observation/visibility conditions forecast to exist due to weather.

4.2 Cover and concealment conditions forecast to exist due to weather.

4.3 Mobility conditions forecast to exist due to weather.

4.4 Effect of extreme conditions of forecast weather on personnel and equipment effectiveness.

V. Risk Factors

As related to mission accomplishment, score each OPFOR course of action in terms of:

5.1 Ability to cope with surprises in terms of U.S. strength or U.S. actions/reactions.

5.2 Freedom from dependence on forces not under own control.

5.3 Freedom from critical dependence on surprise or deception.

5.4 Suitability under unexpected adverse weather conditions.

Figure 1-2

FACTORS IN ENCOA

The situation template is represented by the prior probabilities, which indicate the relative likelihood of the different OPFOR courses of action (i.e., hypotheses) under consideration. The event templates and event matrices are represented by the conditional probabilities, which indicate the relative likelihood that certain events support particular courses of action. The posterior probabilities are represented by the decision support template, which indicates the revised likelihood of the courses of action (i.e., hypotheses) on the basis of observed data. This new estimate of enemy intent is then input to friendly tactical decision making and subsequent action.

Reviews of psychological research in which subjects' final probability estimates have been compared with those prescribed by Bayes' Theorem have shown consistently that humans are suboptimal processors of probabilistic information. Although they typically revise their opinions in the same direction as Bayes' Theorem, they do not revise them enough; they are conservative. This finding could have great implications for IPB, for if analysts using IPB are conservative information processors, then they are not drawing implications from the data as fast as they could be with Bayes' Theorem. Their estimates about enemy courses of action may well be suboptimal because they will not have sufficiently revised their opinions to take full account of the certainty in the data. Consequently, the entire templating process will not convey as much information to commanders as it should. The time available for friendly tactical decision planning and implementation may be reduced considerably if intelligence analysts are conservative information processors.

The potential implications of Bayesian research for IPB are compounded by the fact that intelligence analysts rely on language instead of numerical estimates to convey

uncertain information. It is impossible to directly translate qualitative expressions of uncertainty such as "very likely" into probability values. For while most people would agree that "very likely" means a probability greater than 0.5, there is no general agreement of how much more than 0.5 the probability is (0.8?, 0.9?, 0.99?). The lack of agreement in the use of language to convey probability estimates has been demonstrated with intelligence analysts in a number of experiments.

Research is beginning to identify why individuals are conservative information processors. One reason, for example, is that people tend to search for confirming rather than disconfirming evidence of alternative hypotheses. This is often a suboptimal data collection and revision strategy because data often can confirm many hypotheses at the same time; disconfirming evidence, on the other hand, can quickly eliminate a hypothesis from consideration. Another reason why individuals are conservative information processors is that they tend to use simple rules or judgment heuristics instead of the axioms of probability theory when making subjective probability estimates. As a result of learning why individuals are conservative information processors, researchers have been able to develop training methods and judgment aids to help people make more accurate probability estimates.

General judgment aids, called Probabilistic Information Processing (PIP) systems have been developed to ensure that human judgment is consistent with Bayes' Theorem for a number of intelligence problems (e.g., see Peterson et al., 1976). Such a judgment aid is now being developed to help tactical intelligence analysts revise their judgment about the most likely OPFOR course of action on the basis of new information. The aid will probably operate in the following manner:

(1) the analysts define the n different enemy courses of action (COAs) under consideration (i.e., enter a brief [≤ 10 character title]);

(2) the analysts enter a set of prior probabilities (or odds) for the n potential courses of action (this could be a set of n probabilities summing to 1.0 or a set of n-1 likelihood ratios assessed relative to a selected COA);

(3) for a given datum, the analysts input a brief title and the probability (or odds in terms of a likelihood ratio) of that datum conditional upon each COA being considered;

(4) the analysts inspect the posterior probabilities (or odds);

(5) they revise any posterior probabilities that are counter-intuitive; and

(6) the analysts report on potential enemy COAs based on the probabilities after step 5 or return to step 3 if there are additional data.

The ability to use a Bayesian framework to represent the different judgments inherent in revising the IPB templates is illustrated in the above steps. Step 1 represents the different OPFOR courses of action represented in the situation template. Step 2 represents the relative likelihood of these actions at the end of one iteration of the IPB analysis process. Step 3 represents the judgments in the event templates and event matrices, which indicate the relative likelihood of having collected the newly acquired data on the basis of the enemy actually taking different courses of action. Steps 4, 5, and 6 represent the

intelligence analysts' estimates of enemy intent, as represented graphically in the decision support template.

The aid is being designed so that it can be readily integrated into the proposed iterative IPB analysis process. Consequently, it will provide an important adjunct to the development and utilization of IPB templates. The decision aid is scheduled for initial evaluation during 1981.

2.0 INTELLIGENCE PREPARATION OF THE BATTLEFIELD: DEPENDENCE ON JUDGMENT

"Everyone complains of his memory, no one of his judgment." La Rochefoucauld.

Techniques for gathering intelligence data have been steadily increasing in sophistication, particularly in the areas of electromagnetic and imagery intelligence. In contrast, techniques for interpreting the implications of intelligence data in terms of enemy intent appear to have improved at a more modest rate (e.g., Bowen et al., 1975). Intelligence Preparation of the Battlefield, as described in Circulars TC30-27 (1978) and TC34-3 (in press), represents a major step toward improving procedures for correlating data with enemy intent and/or capabilities.

Intelligence Preparation of the Battlefield (IPB) is an analytical tool to "...help the analyst 'visualize' the variables of how the enemy <u>might</u> fight on a specific piece of terrain at certain times" (TC30-27, 1-5). This is accomplished through "templating," using a series of graphic illustrations of an enemy capability drawn as an overlay over a map. The process whereby input (new intelligence data) is translated into output (templates) necessarily depends on analysts' judgments about the implications of multiple pieces of potentially fallible data. TC30-27 and TC34-3, however, provide little information about how intelligence analysts should make these judgments. This omission may limit the usability and accuracy of IPB as it is proposed in these circulars. The purpose of this paper is to: (1) demonstrate how the execution of IPB procedures depends on human judgment, and (2) indicate how the incorporation of decision-analytic procedures and computerized judgment aids can improve the quality of judgment in IPB.

Over the last twenty-five years, thousands of scientific studies of human judgment and decision making have reached one basic conclusion: human judgment is limited. The conceptualization of human rationality as limited or "bounded," of even flawed, is a relatively new intellectual idea, one initially developed in large part by Herbert Simon, the 1978 Nobel laureate. As the quote by La Rochefoucauld suggests, one's cognitive ability to decide rationally is seldom questioned. In fact, as TC30-27 illustrates, the question may never be raised. People may complain that their decisions were poor because they didn't have enough information or the right information or management information system to keep track of all the information. But they will seldom, if ever, consider that they just didn't have the cognitive or logical tools to make a good decision. Yet scientific research suggests that this is often the case.

Our self-image to the contrary, judgments of experts and lay people alike have been shown to be inaccurate and subject to systematic biases under laboratory and field conditions. We quote Einhorn and Hogarth (1978) on this point:

> "Although the study and cataloguing of judgmental fallibility has had a long history in psychology (see, e.g., Guilford, 1954, chapter 12; Johnson, 1972), an accumulating body of recent research on clinical judgment, decision-making, and probability estimation has documented a substantial lack of ability across both individuals and situations" (Slovic, Fischhoff, & Lichtenstein, 1977; Slovic & Lichtenstein, 1971) p. 395.

The judgment fallibility referred to above is not attributable to motivational effects such as wishful thinking or distortions by expected payoffs. Severe errors of judgment

occur even when participants are encouraged to be accurate and are rewarded for the correct answer.

In light of the scientific findings, judgment/decision researchers have begun developing decision-analytic techniques and computerized aids to help decision makers improve their cognitive tools and skills. (Kibler, Watson, Kelly, and Phelps [1978], and Patterson and Phelps [1980] have developed such procedures to aid Army G-2 and G-3 staff.) These techniques and aids often rely on the principle of divide and conquer. That is, the total problem is divided into a series of structurally related parts. The decision maker is asked to evaluate each alternative under consideration for each of the independent, simpler components that comprise the larger decision problem. Then, computerized decision aids combine all the judgments in order to provide an overall evaluation of each alternative. Many of the aids also provide sensitivity analysis, thereby permitting decision makers to observe the effects of changing their judgments on each alternative. Thus, decision-analytic techniques and computerized aids "conquer" the global decision problem by providing an analytical means of expanding the cognitive limitations of unaided human judgment.

The analysis process proposed for IPB can be improved by incorporating decision-analytic techniques and computerized decision aids. This can be accomplished because the process whereby intelligence data is translated into IPB templates necessarily depends on analysts' judgments about the implication of the information. Scientific research results strongly suggest that decision-analytic techniques and computerized decision aids would improve both (1) the judgmental accuracy of intelligence analysts developing IPB templates, and (2) the communication process for conveying the reasons for these judgments to tactical commanders. By improving the judgmental process inherent within IPB, and

the means for conveying the reasons for these judgments to tactical commanders, the broader IPB analysis process should be improved in turn.

2.1 Conceptual Overview

Figure 2-1 is a simplified representation of the judgmental nature of the intelligence process. The goal of intelligence analysis is to predict enemy intent or capability (right circle in Figure 2-1). Because the enemy intent is unknown (left circle, Figure 2-1), intelligence analysts must base their predictions on information and indicators that serve as cues to determining the unknown enemy intent. Intelligence predictions are accurate to the extent they match the unknown enemy intent as shown in Figure 2-1. Accuracy is thus a function of: (1) how reliably different cues reflect enemy intent and (2) the degree to which analysts base their predictions on the most reliable cues. The identification and integration of these cues to form the intelligence estimate depends totally on the subjective judgment of the intelligence analyst.

The subjective nature of intelligence analysis becomes even more obvious when more realistic representation of the intelligence process, such as those represented in Figures 2-2 and 2-3, are considered. In particular, there are three factors complicating the analyst's judgment shown in Figure 2-2. First, the relationship between the information and indicators is probabilistic; that is, indicators or information correspond to enemy intent less than 100% of the time, causing the indicators to vary in reliability. Second, intelligence analysts may not use some indicators regardless of the reliability. Third, some information, not related to the enemy intent, may influence the analyst's prediction. Analysts must be able to perceive the relative reliability

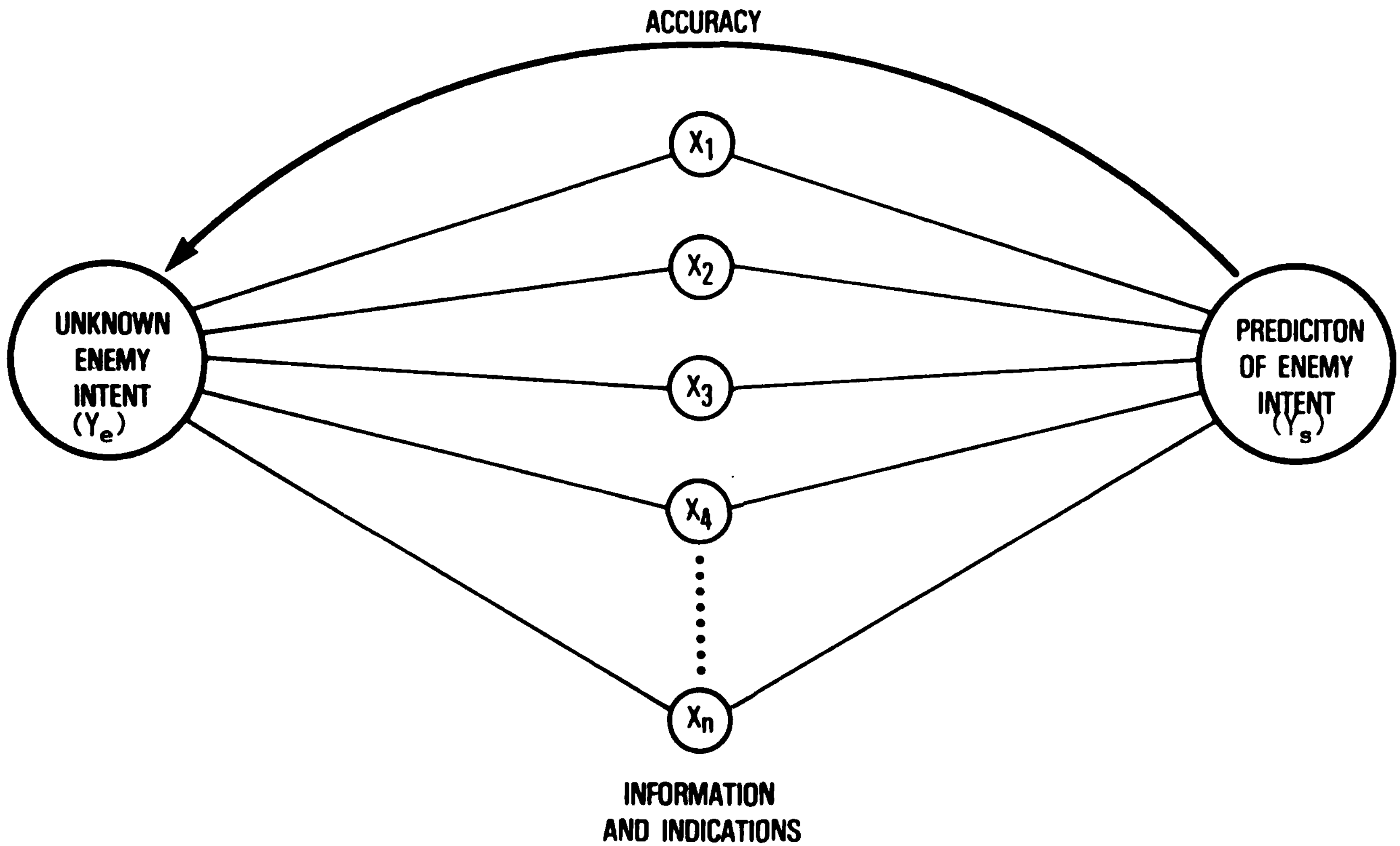

Figure 2-1

PICTORIAL REPRESENTATION OF THE JUDGMENTAL NATURE OF INTELLIGENCE ANALYSIS: AN OVERVIEW

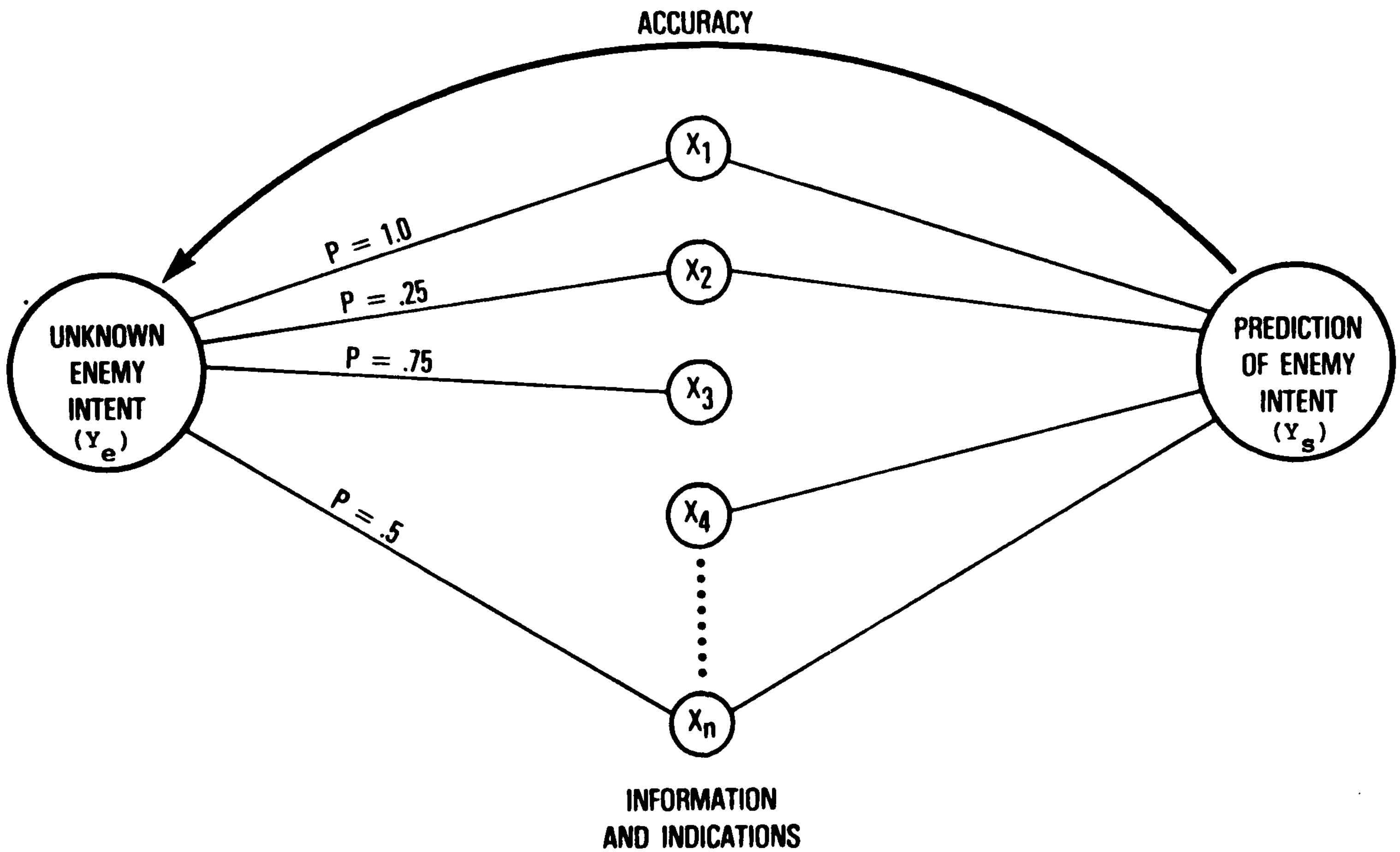

Figure 2-2

PICTORIAL REPRESENTATION OF THE JUDGMENTAL NATURE
OF INTELLIGENCE ANALYSIS: THE UNRELIABILITY OF INFORMATION

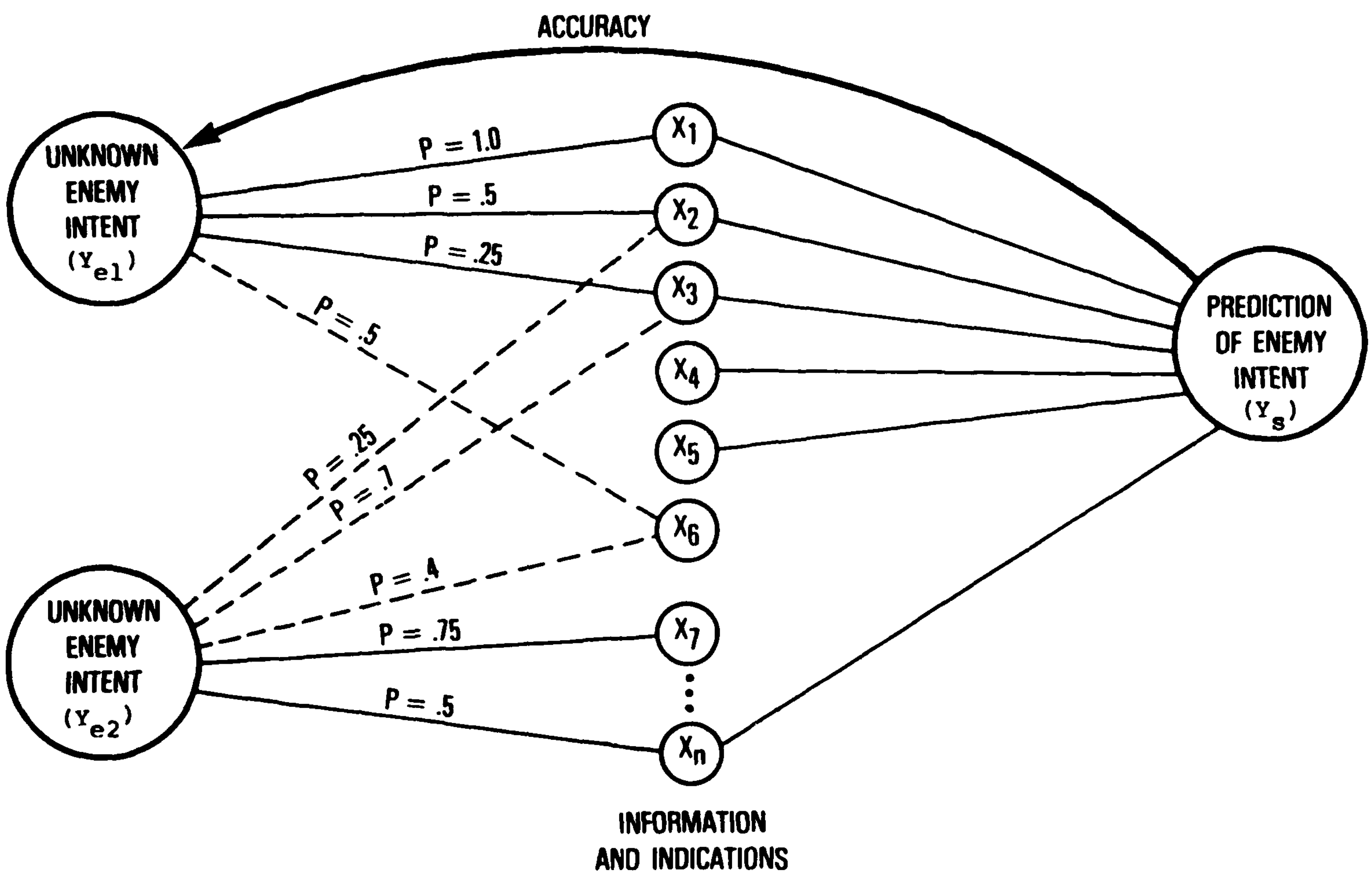

Figure 2-3

PICTORIAL REPRESENTATION OF THE JUDGMENTAL NATURE OF INTELLIGENCE ANALYSIS: MULTIPLE ENEMY INTENT

of indicators and be able to use only the most reliable ones in their analysis, ignoring the unreliable indicators.

At a more conceptual level, the task facing intelligence analysts is one of causal inference. In fact, Figure 2-1 merely represents a minor modification of the "lens model," originally developed by Brunswik (1952, 1956), to describe pictorially the judgmental process of causal inference. In considering the characteristics of causal environments, Brunswik discussed two concepts of particular importance when considering the judgmental nature of intelligence analysis. First, the environment is a causal texture with different events being dependent on each other. As Figure 2-1 implies, cause and effect is conceived of as an "if-then" contingency; $\underline{X}_i$ is clearly dependent on $\underline{Y}_e$, assuming, of course, that $\underline{Y}_e$ occurs first in time. Consequently, a given $\underline{X}_i$ can be considered to be a local representative of (or cue to) the more distal event $\underline{Y}_e$. Second, the dependency between a given cue $\underline{X}_i$ and a given distal state $\underline{Y}_e$, is almost always to some extent equivocal or probabilistic. Hard and fast one-to-one relations between a distal state and cue do not always exist. As a result, a cue can be causally connected with more than one proximal stimulus. The environment therefore ordinarily involves only probable causes and probable effects.

This situation is portrayed by Figure 2-3. The letters Y_{e1} and Y_{e2} indicate that there are two possible causes of the data, or effects, observed by the analyst. The dashed lines indicate the probabilistic nature of some cause-effect relations. For example, X_2 is only related to Y_{e1}; therefore, it is an unequivocal predictor of Y_{e1}. The question facing analysts and commanders is, what data best predict Y_{e1} and what data best predict Y_{e2}?

The proposed IPB analysis process essentially intends to answer this question by templating, that is, by using a series of graphic illustrations of an enemy capability drawn as an overlay for a map. While important, this advance does not go far enough, for it fails to make explicit how input, new intelligence data (X_i), is to be converted into output, templates. This translation process necessarily depends on judgment. TC30-27 and TC34-3 provide little information about how intelligence analysts should make these judgments.

As noted in the introduction, we believe that the proposed IPB analysis process can be strengthened and expanded by incorporating decision-analytic techniques and computerized aids. The next section of this report is directed toward identifying the different types of judgment inherent in each of the five steps in the proposed analysis process. Figures 2-1, 2-2, and 2-3 provide a general conceptual reference for considering judgments of enemy intent.

2.2 The Role of Judgment in Each of the Five IPB Steps

"IPB is a continuous process of analysis and evaluation which is the basis of intelligence operations planning... The purpose of this analysis is to determine and evaluate enemy capabilities, vulnerabilities, and courses of action as the basis for friendly operations planning" (TC30-27, p. iii). To accomplish this purpose, the following five steps in IPB analysis are proposed in TC30-27: (1) mission and threat evaluation, (2) TI zones determination and evaluation, (3) terrain analysis, (4) weather analysis, and (5) threat integration. The next section of this report is directed toward identifying the different types of judgment inherent in each of these five steps, in turn.

2.2.1 Mission and threat evaluation - This is the first step of the proposed IPB analysis process. Mission

evaluation establishes the area of operation and identifies the general battlefield scenario. Threat evaluation of doctrine, tactics, capabilities, and equipment includes a review of the following order of battle intelligence holdings: composition, disposition, strength, tactics, training, logistics, and combat effectiveness. The purpose of this review is to identify the gaps in collected intelligence data and, thereby, guide future collection efforts.

Opposing force (OPFOR) doctrine is to serve as the basis for knowing and understanding the potential adversary and, therefore, implementing mission and threat evaluation. Intelligence analysts are to develop "doctrinal templates" that indicate how the enemy likes to fight. Figure 1 in TC34-3, for example, shows a doctrinal template for a motorized rifle regiment in movement to contact. Intelligence analysts and tactical commanders "seeing" an OPFOR motorized rifle regiment in this disposition, therefore, should conclude that the enemy's intent is to attack.

The doctrinal template thus represents a standard against which order of battle holdings, such as disposition, can be reviewed. If the intelligence team does not know, for the above example, the disposition, composition, or combat effectiveness of the OPFOR motorized rifle regiment in the area of operation, then intelligence data must be collected on the missing order of battle holdings. Once the data are collected, the team can convey their evaluation of OPFOR threat to the commander in the form of a template or otherwise. What Circulars TC30-27 and TC34-3 fail to point out, however, is that the review of order of battle holdings is a complex judgmental process. The intelligence team must evaluate a large amount of incoming data, much of which may be unreliable, in order to decide whether it "knows" the order of battle holdings. In addition, they must integrate

the order of battle holdings, complete or not, into a judgment of enemy intent.

The judgmental nature of order of battle holdings is noted clearly by Bowen, Halpin, Russell, and Staniforth (1975) in Tactical Order of Battle: A State of the Art Survey. Judgment is required because there is simply no alternative to it. We quote Bowen, et al. (1975, p. 45) on this point:

> "Each OB Factor is defined in terms of a number of information elements which indicate the kind of data required to describe the status of that factor in narrative terms. No specific rules exist, either formal or heuristic, for the evaluation of factors or elements, or their combination."

In fact, the lack of "specific rules," or in terms of IPB, "templates," for making Order of Battle judgments is represented in each of the following "key problems" identified by Bowen et al. (1975, pp. 59-60):

"1. There are no standardized methodologies for estimating OB factors....

2. There are no standardized methodologies for incorporating OB factors into the products of OB intelligence.

3. There is no methodology for estimating and reporting reliabilities of OB factor estimates nor the significance of levels of OB factors and their elements relative to general descriptors of the state of enemy forces, such as Combat Effectiveness.

4. There is no realistic and generally accepted definition of the OB factor of Combat Effectiveness. A serious ambiguity of rationale about estimating enemy Combat Effectiveness derives from the point of view of the estimator. (A G2 thinks in terms of the enemy force, while the commander of G3 is concerned with the enemy's net potential effectiveness relative to friendly forces in the existing circumstances.)

5. There is no indication in doctrine of the relative importance of the OB factors, or of their interrelationships.

6. There are no consistent, validated indicators, data aggregates, or data elements for the various OB factors and their elements.

7. There is no methodology for relating the elements of an OB factor to each other, or to the factor itself."

Nor does TC30-27 or TC34-3 offer a "methodology" for making the judgment inherent in threat evaluation. Intelligence analysts are still on their own to combine data on order of battle holdings as they see fit. And, as point #4 above notes, different analysts might combine the data differently. While the proposed IPB analysis may use templates to represent the output of this judgmental process, it provides no means for representing, pictorially or otherwise, the basis for the judgments underlying threat evaluation. In terms of Figure 2-2, for example, TC30-27 and TC34-3 fail to provide an explicit representation of the probable relations between actual enemy intent and the order of battle holdings that represent indicators of intent (X_i); consequently, it offers no explicit rules or heuristics

describing how order of battle holding (X_i) are to be combined into judgments of enemy intent. Scientific research on causal inference indicates that the accuracy of causal inferences can be improved by providing information about probable cause-effect relations represented on the left side of the model in Figure 2-2.

2.2.2 Tactical intelligence zone determination and evaluation - This is the second step in the proposed IPB analysis process. Circular TC30-27 proposes three TI zones to help the commander at the company, battalion/brigade, and corps/division levels, respectively, "... SEE the battlefield" (TC30-27, p. 27). For, according to TC30-27, it is the intelligence analyst's job to make sure that commanders know what the intelligence collection and processing system can and cannot do prior to combat. Toward that end, TI zone determination and evaluation is proposed as a means of telling commanders where to look for targets and indicators of activities that confirm the enemy's adoption of a course of action; what to look for, that is, the critical indicators that must be seen by a certain time; when in the battle sequence to look for them; and what data collection tools to look with. Circular TC34-3 also emphasizes these points, but refers to the TI zones as "areas of interest/influence" (see Chapter 3).

Step 2, therefore, is supposed to identify how the OPFOR threat evaluated in Step 1 should look in the battlefield, in general. Circulars TC30-27 and TC34-3 accomplish this, however, only to a limited degree. There is, for example, no discussion of the general indicators for different OPFOR threats. Yet research by Johnson, Spooner, and Jaarsma (1977) suggests that this would be a valuable addition for tactical commanders. They found that a sample of forty-three captains in the Intelligence Officers Advanced

Course knew only nineteen of the forty-nine separate indications listed in Field Manual 30-5, Combat Intelligence, for the four separate courses of action to be evaluated in Step 1 - attack, defense, delay, and withdrawal.

Just as in the case of threat evaluation, the judgmental conceptualizations shown in Figures 2-1, 2-2, and 2-3 represent the basic judgment problem facing analysts during TI zone evaluation. Doctrinal templates could be developed to help commanders visualize the general indicators representative of different OPFOR threats at different TI zones. In addition, decision aids could be developed to assist in TI zone evaluation. Since individual indications are not perfect predictors of enemy intent, analysts and commanders must know the relative accuracy of individual indicators for each of the different OPFOR courses of action. Furthermore, they must know how to combine individual indicators into a global assessment of threat. TC30-27 provides no indication of how this is to be accomplished in TI zone evaluation. Causal inference research indicates that explicit information about the left-hand side of the causal inference represented in Figures 2-1, 2-2, and 2-3, greatly improves the judmental accuracy of individuals.

2.2.3 Terrain and weather analysis - It is interesting to note that Circulars TC30-27 and TC34-3 do tell analysts how to make judgments about the military aspects of terrain and weather, the third and fourth steps, respectively in the proposed IPB analysis process. The determination of avenues of approach into the commander's TI zone is basic to terrain analysis. To accomplish this, the IPB analysis team must focus its analytical efforts on the following five military aspects of terrain: observation in fields of fire, concealment and cover, obstacles, key terrain, and avenues of approach. Toward this end, the circular contains a factor and subfactor analysis matrix, and a list of terrain requirements.

This information defines the relationships between the physical aspects of terrain, such as surface configuration, vegetation, rock types and soils, and specific military aspects of terrain, such as observation and fields of fire, cover, concealment, and obstacles. Thus, it shows the analysts how to make judgments about the military aspects of terrain. Such analytical efforts facilitate subsequent "mobility analysis" and "line-of-sight analysis," the result of which is a series of map overlays that identify the possible OPFOR avenues of approach and the ..." one avenue of approach (mobility corridor) more favorable than the others" (TC30-27, p. 3-26). They also provide a retraceable procedure for telling the commander how these judgments were arrived at by the analysts.

As in the case of terrain analysis, the IPB analysis process addresses the effects of weather on ground and air mobility and line-of-sight. In a similar fashion, Circulars TC30-27 and TC34-3 address in considerable detail how the "weather factor analysis matrix" and "weather parameter-user matrix" are to be used to relate weather factors to a wide range of military operations. The result of Step 4 is a set of overlays that integrate terrain and weather data in a manner that conveys their <u>possible</u> interactive effects on the capability of forces to move, shoot and communicate within the different mobility corridors in the situation under consideration. The word "possible" is underlined to emphasize the uncertain nature of weather; the best mobility corridor if it's dry can be the worst one if it rains.

2.2.4 <u>Threat integration</u> - This is the fifth and final step in the proposed IPB analysis process. "The objective of threat integration is to relate "how the OPFOR likes to fight" to a specific terrain and weather scenario as the basis for determining "how the enemy might have to fight"

(TC34-3, p. 5-1). This is to be accomplished through the use of situation, event, and decision support templates. These templates are proposed as a means of helping commanders "visualize" enemy capabilities in a particular combat setting. They require intelligence analysts to make a number of different kinds of judgments. Both circulars fail, however, to tell intelligence analysts how to make these judgments. These different judgments are identified below for each template, in turn. In addition, limitations in the circulars are identified, and suggestions are offered for their improvement.

The situation template shows how enemy forces probably would look within the different mobility corridors under consideration. The analysts developing the situation template, therefore, must use the terrain and weather analyses to modify the doctrinal templates for the OPFOR threat evaluated in step #1. Underlying the development of this template is a complex judgmental process. We quote TC30-27 on this point.

> "While the enemy commander may not have unlimited options as to possible courses of action, he will probably have enough options to make the analyst's job of determining probable courses of action extremely difficult. Situation templates are derived based on the best military judgment of the analyst (p. 4-4)."

Unfortunately, TC30-27 provides minimal information on how analysts are to exercise their "... best military judgment ..." when developing the situation templates. In fact, it does not even identify the factors that are to be considered explicitly when making their judgment.

If the situation template is to reflect military judgment about different OPFOR courses of action, it is not

enough to show analysts how to perform a terrain and weather analysis. In addition, Circulars TC30-27 and TC34-3 should identify the factors that are to be considered when making this judgment, for the analysis team will have to incorporate judgments about other, more ambiguous factors, such as perceived U.S. force strength and risk, that an OPFOR commander would certainly consider when selecting a course of action. Furthermore, the circulars should discuss the trade-offs that OPFOR commanders are likely to make when evaluating the utility of different courses of action. Rarely will it be true that one course of action is better than all others on every factor. Enemy commanders will be forced to differentially weight aspects of their doctrine with the characteristics of the situation immediately facing them. Circulars TC30-27 and TC34-3, or supporting documentation, should indicate what these trade-offs are likely to be under different terrain and weather constraints, if the situation template is to accurately represent military judgment under different circumstances. Said differently, it should describe the left side of the lens model in Figure 2-1, which describes how the enemy tends to make judgments of intent.

"The event template is a time and logic sequence of enemy tactical indicators or events which are keyed to a series of situational templates" (TC30-27, pp. 1-6). This is to be accomplished by identifying, on the situation template, "NAMED AREAS OF INTEREST (NAI's)... along each route where the analyst expects to see certain events or activities occur" (TC30-27, p. 4-4). In addition, the analyst is to complete an event analysis matrix that "... correlates WHAT IS EXPECTED (event/activity) to the WHERE and WHEN (Geographical coordinates and time)" (TC30-27, p. 4-6). The event template is to be a combination of the situation template and the event analysis matrix.

Although TC30-27 and TC34-3 provide no example of an event template, it appears that the template must identify the different types of information necessary to confirm or deny the adoption of a particular course of action. To accomplish this, the analysts must make a series of conditional probability judgments analogous to those represented in Figures 2-2 and 2-3. That is, they must say, "If the enemy actually took this particular course of action, then these indicators, and events (e.g., X_2) have a higher probability of being observed than others." The word "probability" must be emphasized because there is not a perfect relationship between indicators, events, and actual enemy intent. The enemy will be expected to use deception measures. In fact, "An integral part of templating is the consideration of deception events associated with each course of action" (TC30-27, p. 4-11). As a result, the intelligence team will be forced to make conditional probability judgments about what indicators and events they think are most indicative of the OPFOR adoption of different courses of action. Neither TC30-27 nor TC34-3, provides any information about the probabilistic relations between different indications and different OPFOR threats, a point made previously when discussing TI zone determination and evaluation.

The third, and final, effort in threat integration is development of a decision support template. This template "... is used to illustrate enemy probable courses of action as the basis for comparing friendly courses of action" (TC30-27, p. 4-13). Described in this way, the decision support template "... is essentially the INTELLIGENCE ESTIMATE in graphic format" (TC34-3, p. 5-10). It represents the analysts' most up-to-date estimate about the relative likelihood of the enemy's potential courses of action. Again, however, TC30-27 and TC34-3 fail to indicate how the judgments underlying their graphic representation are to be made, or in fact are made, by individual analysts. In terms

of Figure 2-1, they fail to describe the right side of the lens model, the reasons for judgments of enemy intent. Psychological research suggests that this may result in decision support templates that (1) are not as accurate as they could be, and (2) do not facilitate communication between analysts and commanders as much as they could otherwise.

2.2.5 Template Revision - Up to this point, there has been no discussion of the dynamic nature of the proposed IPB templating process. The templating process is not supposed to stop with the first decision-support template. "They are to be changed, deleted, or redone as conditions and situations demand. Used properly, they provide for continuing integrative analysis of OPFOR capabilities, vulnerabilities and courses of action" (TC30-27, p. 1-5). This iterative revision process is critical to friendly tactical decision making. Friendly commanders will make particular decisions, and take subsequent actions, on the basis of intelligence analysts' estimate of enemy intent. This estimate will be revised continuously upon the arrival of new information. The faster the information can be correlated with enemy intent, the more time friendly commanders will have for tactical decision making and action.

Template revision is represented pictorially in Figure 2-4. Notice that the decision support template at the end of one cycle of the process is the situation template at the beginning of the next cycle. The event templates and event analysis matrices are used to convey, as quickly as possible, revised estimates about the likelihood of different OPFOR courses of action. These revised estimates of enemy intent are represented in the decision support template.

The dynamic nature of the proposed templating process is heavily dependent on a complex judgment process.

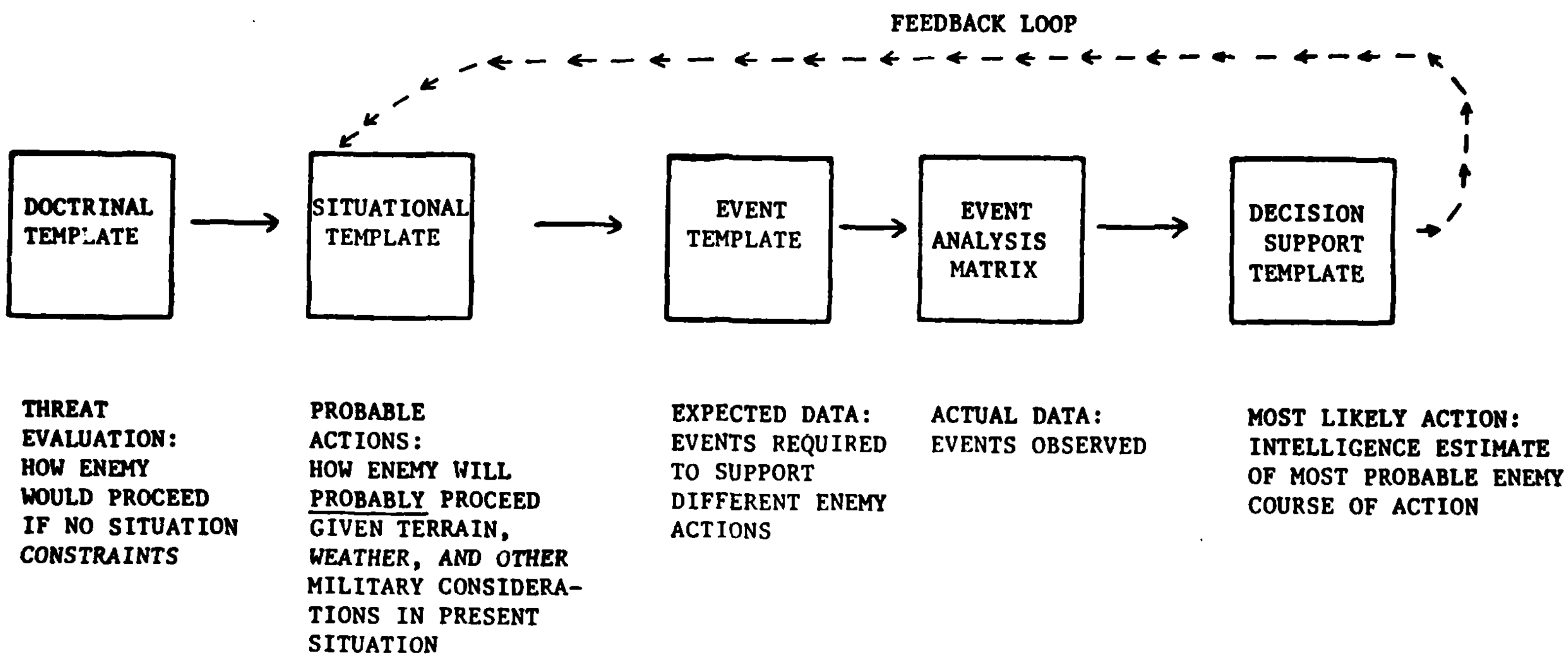

Figure 2-4

SCHEMATIC REPRESENTATION OF THE DYNAMIC NATURE OF THE PROPOSED IPB TEMPLATING PROCESS

First, intelligence analysts using the proposed IPB analysis process must identify the initial hypotheses regarding possible enemy courses of action and, if possible, an initial estimate of the most probable course of action. As discussed above, this requires consideration of how the enemy generally makes judgments of intent, in addition to an evaluation of how physical terrain and weather factors favor different courses of action.

Second, intelligence analysts using the proposed IPB process must indicate the events that are likely to be observed for each course of action. The information within each event template represents, in qualitative or intuitive terms, conditional probabilities. If the enemy actually took a particular course of action, then certain events have a higher probability of being observed than others. The information within all the event templates represents, again in qualitative terms, likelihood ratios that indicate how much more likely certain events are to be observed for one course of action than another. While these probabilistic judgments are not being made quantitatively when analysts develop the event templates, they are being made intuitively in order to guide data collection in an uncertain environment.

Third, the analysts must use the many pieces of potentially fallible data reported in the event matrices (and other sources) to revise their initial hypotheses about the enemy's most likely course of action. The data reported in the event matrices also represent conditional probabilities, for certain events have a higher probability of being observed <u>only</u> when the enemy is actually taking a particular course of action. As a result, subsequent decision support templates represent, again in qualitative terms, the analysts' most up-to-date estimate about enemy intent, as revised by newly acquired intelligence data.

The three areas of judgment inherent in the revision process can be represented conceptually by the modified lens model shown in Figure 2-5. The situation template indicates the analysts' military judgment about the likelihood of different enemy courses of action. The event templates indicate events the analysts expect to see in support of each course of action. The event analysis matrices indicate the events subsequently reported in intelligence data. The events "seen" are matched with those "expected" in order to develop a revised estimate of enemy intent, which is represented in the decision support template. The decision support template is now the situation template for the next iteration.

There are different ways to represent quantitatively the three areas of judgment inherent in the template revision process. These procedures will be considered in some detail in the next section when we consider different decision aids for supplementing, and thereby improving the proposed IPB analysis process. Suffice it to say for now, that these procedures often rely on the principle of divide and conquer. First, the total problem is divided into a series of structurally related parts. The intelligence analyst is asked to evaluate each OPFOR alternative under consideration for each of the independent, simpler components that comprise the larger intelligence problem. Then, computerized judgment aids combine all the judgments in order to provide an overall evaluation of each alternative. Many of the aids also provide sensitivity analysis, thereby permitting intelligence analysts to observe the effects of changing their judgments on the overall score for each alternative. Thus, decision-analytic techniques and computerized aids "conquer" the global decision problem by providing an analytical means of expanding the intelligence analyst's cognitive skills.

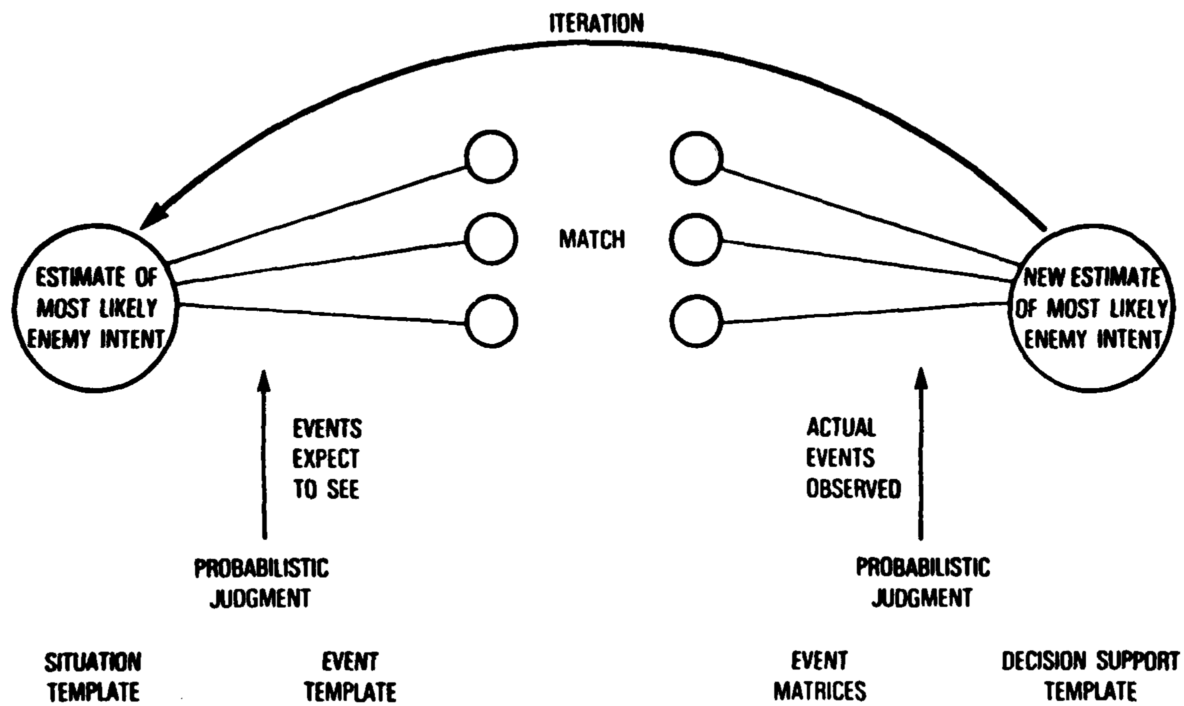

Figure 2-5

LENS MODEL REPRESENTATION OF JUDGMENTS INHERENT IN REVISION OF IPB TEMPLATES

2.3 Summary

The purpose of IPB is to improve procedures for correlating data with enemy intent. This purpose is to be accomplished by templating, an analytical tool designed to help analysts and commanders visualize the battlefield. While important, this advance does not go far enough, for it fails to make explicit how input, new intelligence data (X_i), is to be converted into output, templates. This translation process necessarily depends on judgment. However, Circulars TC30-27 and TC34-3 provide practically no information about how these judgments are to be made by analysts.

Section 2.0 described the necessary dependence of IPB on human judgment. In particular, it provided a conceptual overview of the causal inference problem inherent in intelligence analysis, and identified the different types of judgment inherent in each of the five steps in the IPB analysis process proposed in Circulars TC30-27 and TC34-3. The general conclusion was that except for terrain analysis (Step 3) and weather analysis (Step 4), these circulars fail to tell intelligence analysts how they are to make the judgments necessary to implement the proposed IPB analysis process. The lack of, and need for, procedures for helping analysts make the judgments in the other three steps of the proposed IPB analysis process were discussed at length.

It is important to note that the proposed IPB analysis process is a dynamic one, for the decision-support template at the end of one cycle of the process is the situation template at the beginning of the next cycle. Three areas of judgment are inherent in the revision process. These areas of judgment were represented conceptually within the causal inference model presented in Section 2.0. Analytical procedures and computerized judgment aids for assisting analysts tasked with making these, and the other judgments in the proposed IPB analysis process, are described in Section 3.0.

3.0 CAUSAL INFERENCE RESEARCH FINDINGS AND COMPUTER-BASED JUDGMENT AIDS APPLICABLE TO INTELLIGENCE PREPARATION OF THE BATTLEFIELD

The previous section of this report described the different types of judgments inherent in the IPB analysis process, as proposed in TC30-27 and TC34-3. This section describes the results of research studying how well people make these judgments. The research strongly suggests that people's ability to make these judgments can be improved substantially by giving them access to computerized decision aids. These aids, therefore, represent an important means of improving the proposed IPB analysis process.

This section is divided into two parts. The first part describes causal inference research findings and computerized judgment aids particularly relevant to the development of the initial decision support template. The second part describes research and aids particularly relevant to the revision of decision support templates on the basis of newly collected data represented in the event template and event analysis matrix. The difference in emphasis is used here to indicate the strengths of each class of judgment aids; however, both types of aids can be used in initial development and subsequent revision of decision support templates.

3.1 Research and Aids Applicable to the Development of the Initial Decision Support Template

As discussed in Section 2.0, the decision support template must identify the different possible enemy courses of action and the most likely enemy choice. To make this latter judgment, the analysis team must integrate individual judgments of OPFOR doctrine with those pertaining to the

military aspects of terrain and weather. In addition, the analysis team must (1) incorporate judgments about other, more ambiguous factors (or attributes), such as perceived U.S. force strength and risk, that an OPFOR commander would certainly consider when selecting a course of action, and (2) estimate the trade-offs that OPFOR commanders are likely to make when evaluating the utility of different courses of action. For rarely will it be true that one course of action is better than all others on every value-relevant attribute. Enemy commanders will be forced to differentially weight aspects of their doctrine with the characteristics of the situation immediately facing them. By accurately predicting these trade-offs, and by representing them in the first decision support template, intelligence analysts can provide friendly commanders with time for efficient tactical decision planning.

This section reviews laboratory research on causal inference conducted within the Brunswickian framework represented in Figures 2-1, 2-2, and 2-3. In particular, Section 3.1.1 reviews research investigating the accuracy of the judgments of individual participants and the obtained level of interpersonal agreement in different settings where participants <u>had</u> <u>no</u> <u>access</u> to judgment aids. Section 3.1.2 reviews research findings regarding the obtained levels of individual accuracy and interpersonal agreement when participants <u>had</u> <u>access</u> to judgment aids. Section 3.1.3 describes a judgment aid designed to assist Staff Intelligence Officers (G2/S2s) in looking at the battlefield from the perspective of enemy commanders <u>prior</u> to their decision to adopt a course of action; therefore, it provides an aid to the development of the initial decision support template.

3.1.1 <u>Research findings: without judgment aids</u> - Considerable psychological research on the accuracy of expert judgment and the level of interpersonal agreement has

been conducted within the conceptual framework of Brunswik's Lens Model (1952, 1956). This research has been conducted, for example, with clinical psychologists (e.g., Goldberg, 1970; Hammond, Hursch, and Todd, 1964; Meehl, 1954), stockbrokers (Ebert and Kruse, 1978; Slovic, Fleissner, and Bauman, 1972), radiologists (Hoffman, Slovic, and Rorer, 1968), histologists (Einhorn, 1974) and highway engineers (Adelman and Mumpower, 1979). In situations with criterion data, the judgmental accuracy of these experts often has been found to be lower than expected by the researchers. In addition, interpersonal agreement has been found to be surprisingly low.

Research by Ebbesen and Konecni (1975) with jurists and Phelps and Shanteau (1977) with livestock experts indicates that the level of judgmental accuracy and interpersonal agreement can be a direct function of the characteristics of the judgmental task facing the experts. For example, both found that the degree of correlation between attributes will affect the number of attributes apparently used by experts.[1] This general conclusion is further supported by controlled laboratory research which has shown that (1) judgmental accuracy and agreement can be increased or decreased by manipulating characteristics of the task and (2) that relatively high levels of accuracy and agreement can be maintained under conditions that normally prevent it by using judgment aids that provide persons with information about task characteristics.

---

[1]The word "apparently" is used because no statistical analysis procedure can identify how a person actually used the information; all it can do is identify the model of information use that best predicts the person's judgment. See Hoffman (1960) for an early, yet detailed discussion of this point.

Causal inference tasks can be characterized in terms of their formal characteristics. Formal task characteristics refer to the statistical relationships among task variables. Research on the following two task characteristics is of particular relevance to effective intelligence analysis: (a) the predictability of individual pieces of information (called cues or attributes), and (b) the total predictability of all the cues in the inference task.[2]

In the standard procedure for studying causal inference, the participant is shown a set of "cases" on which to make judgments. In terms of Figure 2-1, each case is described by values on a set of cues (or attributes) (X_i) that participants use to make a prediction (Y_s) of the criterion or "cause" (Y_e). Formal task characteristics determine the relationship between the criterion and the cue values. Participants are shown the correct answer after each case, so they can learn how to make accurate causal inferences. After learning has occurred, judgmental accuracy (or "achievement") is measured statistically.

Hursch, Hammond, and Hursch (1964) and Tucker (1964) have shown that achievement can be partitioned into statistically independent components. Under conditions in which the criterion values can be expressed as a linear function of the attribute values, Tucker's formula can be stated as follows:

$$\underline{r}_{\underline{a}} = \underline{G}\underline{R}_{\underline{s}}\underline{R}_{\underline{e}} \qquad (1)$$

---

[2]The words "cue" and "attribute" are used interchangeably throughout this section to provide continuity between researchers within the Brunswikian framework who use the word "cue," and decision analysts developing judgment aids who use the word "attribute." The word "cue" is more appropriate in this section because information (X_i) is a "cue" to the correct criterion (Y_e), not an attribute of it.

In this formulation, $\underline{r}_a$ is the correlation between criterion values and the participant's judgments of these values; hence, $\underline{r}_a$ provides a measure of achievement. $\underline{G}$ is the correlation between the linear variance in the task system and the linear variance in the participant's judgment system. Since a high $\underline{G}$ indicates that the properties of the participant's judgment system match the properties of the task system, $\underline{G}$ provides a measure of the participant's knowledge of task properties. $\underline{R}_s$ is the linear multiple correlation between the attribute values and the participant's judgments. Since a high $\underline{R}_s$ indicates that participants are making consistent judgments based on whatever (not necessarily correct) knowledge of the task they have acquired, it provides a measure of "cognitive control" (Hammond and Summers, 1972) or the ability of participants to consistently make judgments in the way they want to do so. $\underline{R}_e$ is the linear multiple correlation between the attribute and criterion values; it is the task's counterpart of $\underline{R}_s$ and, therefore, is a measure of task consistency or predictability.

It is important to note that $\underline{R}_s$ is statistically independent of $\underline{G}$. This indicates that even if $\underline{G}$ reaches unity (indicating perfect knowledge), if $\underline{R}_s$ were less than unity (indicating imperfect control), achievement ($\underline{r}_a$) would be less than the level of task predictability ($\underline{R}_e$). Conversely, the same results could occur if $\underline{R}_s$ equalled 1.0, but $\underline{G}$ were less than 1.0. Therefore, two persons can have identical achievement indices ($\underline{r}_a$) for different reasons: one because of perfect knowledge ($\underline{G}$ = 1.0) but imperfect control ($\underline{R}_s$ < 1.0), the other because of perfect control ($\underline{R}_s$ = 1.0) but imperfect knowledge ($\underline{G}$ < 1.0) This distinction is critical when discussing the effects of formal task characteristics on the accuracy of causal inference.

Variation in the predictability of individual cues, for example, affects achievement by affecting both the

acquisition and the utilization of knowledge about the causal environment. For example, Uhl (1963) investigated the effects of varying the predictability of individual cue validity by constructing 7 three-cue tasks ranging in individual cue predictability from one extreme where all the cues were equally predictable (i.e., $\underline{r}_{e1} = \underline{r}_{e2} = \underline{r}_{e3}$, where $\underline{r}_{ei}$ is the correlation between the cue and the criterion), to the extreme where only one cue predicted the criterion. In all tasks the cues were interval in nature, linear and orthogonal. And in all cases, overall task predictability was unity (i.e., $\underline{R}_e = 1.0$). In these tasks, achievement ($\underline{r}_a$) was poorest in tasks where individual cues had equal predictability. Lower achievement in such tasks resulted from both lower knowledge ($\underline{G}$) and lower cognitive control ($\underline{R}_s$) than in the tasks with a larger disparity in the predictability of individual attributes. Dudycha and Naylor (1966) found similar results. <u>These results are important to intelligence analysts because, according to the report by Bowen, et al. (1975) reviewing Tactical Order of Battle, intelligence analysts have to consider many information elements of relatively equal importance</u>.

Variation in overall task predictability ($\underline{R}_e$) affects achievement ($\underline{r}_a$) by leading to variation in cognitive control ($\underline{R}_s$), not knowledge ($\underline{G}$). For example, Uhl (1966) investigated the effects of task predictability on achievement by constructing four three-attribute tasks ranging in predictabilities of 1.0, 0.67, 0.33, and 0.00. The attributes were orthogonal and linear; individual attribute predictability ($\underline{r}_{ei}$) was 0.775, -0.225, and 0.000, respectively. Uhl (1966) found that there was a positive linear function between task achievement and task predictability; i.e., $\underline{r}_a$ decreased as $\underline{R}_e^2$ decreased. The cause of lower achievement, however, was not lower knowledge ($\underline{G}$), but lower cognitive control ($\underline{R}_s$), for there was a positive linear function between task predictability and cognitive

control; i.e., as $\underline{R}_e^2$ decreased, $\underline{R}_s^2$ decreased. Furthermore, the match between $\underline{R}_e^2$ and $\underline{R}_s^2$ was quite high, except for $\underline{R}_e^2$ = 0.00; specifically, $\bar{\underline{R}}_s^2$ was 0.90, 0.66, 0.39, and 0.29 for $\bar{\underline{R}}_e^2$ of 1.0, 0.67, 0.33, and 0.00, respectively. Therefore, persons not only matched the predictability of individual attributes, they matched overall task predictability (or inversely, randomness) as well. Dudycha and Naylor (1966) and Schmitt, Coyle, and King (1976) found similar results. This finding is of particular importance to intelligence analysts because the inherently low level of overall predictability in some intelligence tasks suggests that analysts may perform suboptimally because of cognitive inability to utilize their knowledge of enemy activities in a highly consistent fashion.

Controlled laboratory research has shown that formal task characteristics also affect interpersonal agreement (see Brehmer, 1976, for an extensive review of this literature). The procedure for this research has two stages: a training stage in which each participant learns to make causal inferences in a certain way, and a conflict stage in which participants make judgments jointly. The lens model equation is again used, but now $\underline{r}_a$ is the correlation between the judgments of the two participants; hence, $\underline{r}_a$ represents the overall level of agreement between the participants, $\underline{G}$ represents the similarity in their judgmental policies (i.e., the way they make causal inferences), and $\underline{R}_{s1}$ and $\underline{R}_{s2}$ represent their levels of judgmental consistency.

Two task characteristics affecting interpersonal agreement that are of particular relevance to intelligence analysis are overall task predictability and cue intercorrelations. Regarding the former, the lower the level of task predictability, the lower the level of interpersonal understanding and agreement. The cause of this result is lower

judgmental consistency (or cognitive control) on the part of each participant, not policy similarity. This occurs because people try to predict the randomness (i.e., uncertainty) in the causal inference task. As a result, they make it difficult for the other person to understand how they make their judgments, and thereby, how to resolve the disagreement. In fact, Brehmer (1972) found that low levels of task predictability could even cause people who began the conflict stage in perfect agreement to finish it in disagreement because of judgmental inconsistency. Therefore, low overall task predictability, the same formal task property that causes suboptimal judgmental accuracy, also causes interpersonal disagreements. Research indicates that computerized judgment aids improve the levels of both accuracy and agreement.

Cue intercorrelations also introduce certain constraints on the relationship among judgmental accuracy, interpersonal agreement, and policy similarity. If the cues have a high positive correlation, for example, then a person can rely totally on an invalid cue and still be accurate because it does not matter which cue is used to predict the criterion. Furthermore, that person's judgments will agree with those of another person who is, in fact, relying on the valid cue. If the situation should change, however, such that the cues no longer have a high positive correlation, then the two persons will be in strong diagreement. Furthermore, they will have great difficulty resolving their dispute because both will point to their recent successes, although only one of the people actually knows which cue is the valid indicator.

3.1.2 Research findings: with judgment aids - One question given considerable consideration by researchers studying causal inference research has been, how can achievement or accuracy be improved? The traditional answer, consistent with a stimulus-response-reinforcement orientation,

has been to provide the learner with outcome feedback, that is, the correct answer after each learning trial. This is conceptually the same answer given when intelligence analysts are forced to "learn by experience." Yet, as Todd and Hammond (1965, pp. 429-430) point out,

"Outcome feedback may be appropriate when the learning task is sufficiently simple that the subject can associate specific responses with specific stimulus configurations, . . . [but] few psychologists would seriously entertain the idea that such specific associations are what is being learned in multiple-cue probability tasks in which very large numbers of such associations are typically required. More probably, it is the relations between cues and criteria which are learned. Outcome feedback, however, does not appear to be appropriate for learning probability relations, because it yields information which is restricted to a comparison of end results--the comparison of the response with the correct answer. . . . Feedback which directly informs the S about the relation of his cue utilizations (r_{si}) should make it possible for the S to adjust his cue utilizations in the direction of the cue validities."

Feedback that permits persons to compare the formal properties of their judgmental system (right-hand side of Figure 2-1) with those of the task system (left-hand side of Figure 2-1) has been called "lens model feedback" (Todd and Hammond, 1965, p. 431), or more recently, "cognitive feedback" (Hammond, Stewart, Brehmer, and Steinmann, 1975, p. 293). Computer graphics devices have been used widely as decision aids for presenting cognitive feedback. Hammond (1971), for example, has shown how computer graphics can be used to provide participants with a pictorial and quantitative representation of achievement in terms of the lens model and lens model equation, respectively. Figure 3-1 provides an example of such a pictorial, quantitative

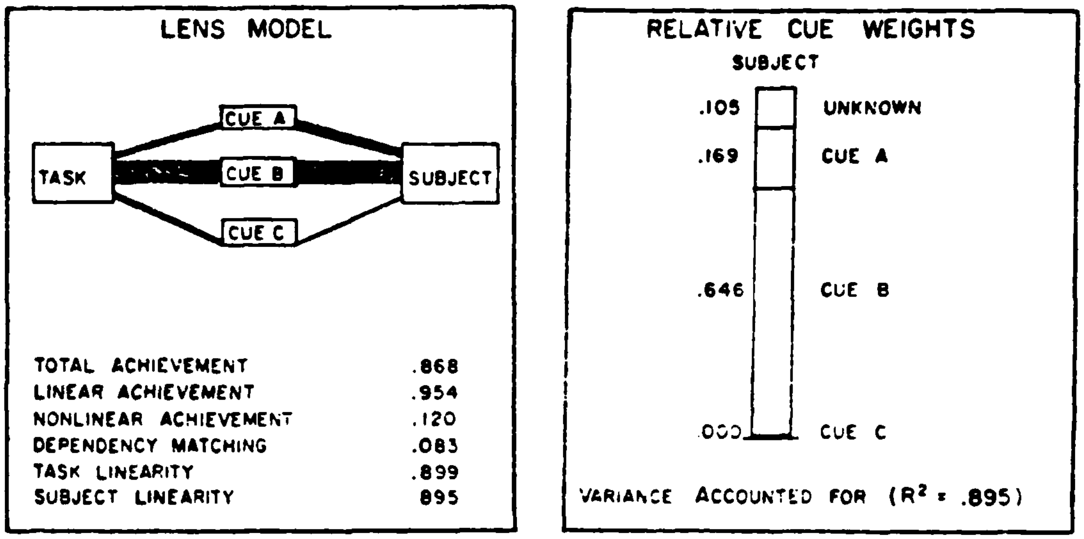

Figure 3-1

EXAMPLES OF COGNITIVE FEEDBACK DISPLAYS
(From Hammond and Boyle, 1971)

display. Cognitive feedback provided by such displays has been shown to lead to both a faster rate of learning and a higher level of achievement ($\underline{r}_a$) than outcome feedback in tasks that varied the (a) disparity of the individual predictabilities of the attributes (Lindell, 1976), and (b) the overall level of task predictability (Schmitt, Coyle, and King, 1976). While this finding is due in some part to greater knowledge acquisition ($\underline{G}$) with cognitive than outcome feedback, it is primarily due to higher cognitive control with cognitive feedback. <u>This finding strongly supports our position that computerized judgment aids will improve the skills and judgmental accuracy of experienced intelligence analysts using IPB by greatly increasing their ability to implement their knowledge</u>.

Laboratory research by Hammond and Brehmer (1973) and by Summers, Taliaferro, and Fletcher (1970a) has found that computerized judgment aids that quantitatively describe how people combine information on multiple cues into an overall judgment greatly enhances interpersonal understanding. The reason for this is that decision aids overcome the inaccuracy and inconsistency of verbal self-reports. Verbal reports are often inaccurate because people inaccurately estimate the weight they place on various attributes. Research by Cook and Stewart (1975) and Summers, Taliaferro, and Fletcher (1970b) indicates that people sometimes underestimate the weight they place on important factors and overestimate the weight they place on unimportant factors when compared to a quantitative analysis of their judgments. To put it bluntly, people are sometimes unaware of their own judgment policy; therefore, it is not surprising that they inaccurately describe it.

Decision-analytic techniques and computerized judgment aids can be used to describe explicitly (1) how different persons make their causal inference judgments,

(2) the intercorrelations perceived by different persons, and (3) how attribute intercorrelations can affect judgmental accuracy and interpersonal agreement. Such capabilities, therefore, provide an explicit means of overcoming formal task properties that research (e.g., Brehmer, 1974; Hammond and Mumpower, 1974) indicates can cause interpersonal disagreements. Furthermore, such capabilities have been used to resolve actual conflicts. (See Hammond, Rohrbaugh, Mumpower, and Adelman, 1977, for a review.) For example, Hammond and Adelman (1976) helped the Denver City Council resolve a conflict over the type of handgun bullet to be used by the Denver Police Department by showing that there was <u>not</u> a perfect relationship between a bullet's stopping effectiveness and its severity, as was perceived by many parties to the conflict. Hammond, Rohrbaugh, Mumpower, and Adelman (1977), and Kelly (1979) provide reviews describing how judgment aids have been used for different substantive problems.

3.1.3 <u>A judgment aid for IPB</u> - Multi-attribute utility assessment (MAUA) techniques and judgment aids can be used to help intelligence analysts develop the initial decision support template. They provide (1) a logically defensible conceptual structure for structuring the factors used in estimating enemy intent, (2) techniques for estimating the necessary trade-offs inherent in estimating enemy intent, (3) an analytical procedure for combining the multiple trade-off judgments into an overall assessment of the likelihood of different enemy courses of actions, (4) a means of systematically investigating the implications of differences of opinion in judgment, and (5) a means of conveying all this information pictorially, thus providing an important adjunct to the proposed pictorial format of the decision support template. As a result, a computerized judgment aid called ENCOA (Enemy Courses of Action) has been developed to assist Staff Intelligence Officers (G2/S2s) develop the

initial decision support template; see Patterson and Phelps (1980) for a complete description.

The utilization of ENCOA, like all MAUA-based decision aids, requires five basic steps. First one develops a hierarchy of attributes (or criteria) that structure the process of evaluating the different alternatives. Second, one scores each of the alternatives on each of the attributes at the bottom of the hierarchy. Third, one specifies the relative importance (or weight) of the multiple attributes within each level of the hierarchy. Fourth, one combines the scores (Step #2) and relative weights (Step #3) to obtain an overall value for each alternative. Fifth, one performs sensitivity analyses to determine what conditions will change the conclusion of the MAUA effort. Each step is considered below at both a general level and for ENCOA in particular.

A MAUA model is hierarchical in nature, starting with the specified top-level factor for which an overall score is desired. This factor is successively decomposed into subfactors in descending levels of the hierarchy such that each successive level is more specific than the one preceding. At the lowest level of the hierarchy are predictable or highly observable characteristics of the alternatives under evaluation.

Figure 3-2 shows the MAUA hierarchy in FNCOA for assessing the overall utility (or value) of the different courses of action open to OPFOR commanders. There are five major factors for evaluating the overall utility of each alternative: Terrain, U.S. Forces, OPFOR Forces, Weather, and Risk. Each of these higher level factors is decomposed into more observable subfactors. Figure 3-3 provides the definitions for these lower level subfactors. These subfactors are defined in a manner that permits intelligence analysts

	FACTOR CATEGORY	FACTOR
0. Mission Accomplishment	1.0 Terrain Factors	1.1 Fields of Fire
		1.2 Cover and Concealment
		1.3 Mobility
		1.4 Seize/Deny Key Terrain
		1.5 Observation
		1.6 Natural/Artificial Obstacles
	2.0 U.S. Force Factors	2.1 Disposition
		2.2 Strength and Condition
		2.3 Reserves
		2.4 Logistic Support
		2.5 Probable Actions/Reactions
		2.6 Command and Control
	3.0 OPFOR Force Factors	3.1 Current Disposition
		3.2 Strength and Condition
		3.3 Reserves
		3.4 Logistic Support
		3.5 Command and Control
	4.0 Weather Factors	4.1 Observation/Visibility
		4.2 Cover and Concealment
		4.3 Mobility
		4.4 Extreme Weather Effects
	5.0 Risk Factors	5.1 U.S. Actions/Reactions
		5.2 Dependence on Other Commands
		5.3 Dependence on Surprise/Deception
		5.4 Unexpected Weather

Figure 3-2

MULTI-ATTRIBUTE UTILITY STRUCTURE FOR
ASSESSING THE OVERALL VALUE OF
DIFFERENT COURSES OF ACTION

I. Terrain Factors

As related to mission accomplishment and considering current OPFOR doctrine, score each OPFOR course of action in terms of how well it:

1.1 Exploits field of fire afforded by terrain features.

1.2 Exploits cover and concealment afforded by terrain features.

1.3 Exploits mobility provisions due to terrain features.

1.4 Accomplishes rapid seizure or denial of key terrain.

1.5 Exploits observation provisions of terrain.

1.6 Exploits or accommodates natural and artificial obstacles.

II. U.S. Force Factors

As related to mission accomplishment and considering current U.S. doctrine, score each OPFOR course of action in terms of how well it exploits what you know or estimate about:

2.1 U.S. disposition.

2.2 U.S. strength and condition.

2.3 U.S. reserves.

2.4 U.S. logistic support.

2.5 Probable U.S. actions/reactions.

2.6 U.S. command and control capabilities/vulnerabilities.

III. Opposing Force Factors

As related to mission accomplishment and considering current OPFOR doctrine, score each OPFOR course of action in terms of how well it exploits or accommodates:

3.1 OPFOR current disposition.

3.2 OPFOR strength and condition.

3.3 OPFOR reserves.

3.4 OPFOR logistic support.

3.5 OPFOR command and control capabilities/vulnerabilities.

IV. Weather Factors

As related to mission accomplishment, score each OPFOR course of action in terms of how well it exploits:

4.1 Observation/visibility conditions forecast to exist due to weather.

4.2 Cover and concealment conditions forecast to exist due to weather.

4.3 Mobility conditions forecast to exist due to weather.

4.4 Effect of extreme conditions of forecast weather on personnel and equipment effectiveness.

V. Risk Factors

As related to mission accomplishment, score each OPFOR course of action in terms of:

5.1 Ability to cope with surprises in terms of U.S. strength or U.S. actions/reactions

5.2 Freedom from dependence on forces not under own control.

5.3 Freedom from critical dependence on surprise or deception.

5.4 Suitability under unexpected adverse weather conditions.

Figure 3-3

DEFINITION OF ENCOA SUBFACTORS

to score each alternative on each subfactor, the second step in using ENCOA.

ENCOA uses a relative scoring system. The best OPFOR course of action on each subfactor is given a score of 100, the worst OPFOR course of action on that subfactor is given a 0, and other courses of action are given intermediate values between 100 and 0 relative to the best and worst alternatives, respectively. A relative scoring system is used instead of an absolute scoring system because of the difficulty, if not the impossibility, of trying to define a true zero level on each subfactor. An absolute scoring system necessitates defining a true zero level of performance and then scoring systems proportional to how far they exceed that zero level; a "relative" scoring system arbitrarily selects the least desirable outcome on each criterion as a relative zero, and then scores each of the other systems proportional to the magnitude of the difference between that system and the one with the lowest score. A relative scoring system is used in ENCOA because of the difficulty in defining a true zero level for the attributes. As a result, some caution is required in interpreting the numerical scores, for while one can make relative comparisons (e.g., better or worse), one cannot make absolute comparisons (e.g., good or bad).

The third step in assessing total utility is to specify the relative importance (or weight) of the attributes within the hierarchy. The factors on each level are compared, by proceeding from bottom to top, to determine the relative importance of the range of variation across the options. Since the options are scored on a relative scale, the weights represent the importance of the range between the best and worst options on each attribute at that level of the hierarchy. For example, if the difference between the OPFOR courses of action on fields of fire is the most important

difference among any of the terrain subfactors, it is arbitrarily given a weight of 100. The intelligence analyst then must assign a relative weight to the second most important difference among the terrain subfactors, the third most important, and so on, each weight reflecting the percentage worth of the difference under consideration to that judged most important and assigned a weight of 100. After weighting all the terrain subfactors, the analysts weight all the subfactors within each of the other four major factors or categories.

After all the above scores and weights have been entered by the user, the final judgmental operation required is to assign importance weights to each of the five categories. To accomplish this, a set of five factors consisting of the top-weighted factor from each category is presented. Again, the intelligence analyst is required to judge the relative importance of the magnitude of the difference between the courses of action scored poorest and best on each of these factors and to rank and weight these differences just as before. This operation has the effect of adjusting factor weights by the importance weighting of the category of which each is a part. If, for example, the four factors under weather had been assigned weights of 100, 80, 50, and 30 on the basis of their relative importance within the weather category, and the weight then assigned the weather category was 50, the adjusted weights for the weather factors would become 50, 40, 25, and 15.

As Kibler et al. (1978) discuss at length, the process of assigning scores to courses of action (i.e., specifying the one which is best on a factor, worst, and intermediate) is found by most people to be an easy task. The process of assigning importance weights to magnitudes of difference is, however, initially an unfamiliar way of thinking for most. Unless users of the decision aid are

carefully briefed and prompted in early trials, there is a tendency for many to slip into the conventional (but, in this context, erroneous) pattern of assigning weights to the generally perceived importance of a factor, rather than to the importance of the difference in value between the extreme courses of action on a factor. With practice, the correct frame of reference for judging importance weights becomes routine. Even then, however, there remains an uncertainty band around the importance weights entered into the model. The individual may have entered a weight of, say, 70 for a factor but was really uncertain as to whether that value might just as well have been 60 or 80. When such uncertainty enters the picture (as is usually the case), it is of great importance to know whether variation of the judgmental inputs within the decision maker's band of error would shift the indicated course of action selection from one option to another--a matter of sensitivity testing.

In the fourth step, one combines the scores of step two and the relative weights of step three to obtain an overall utility value for each alternative. The utility values are obtained by multiplying the bottom-level value score of each option by its relative weight to obtain a score for the option at the next level. This procedure continues until each option receives an overall utility score at the top of the hierarchy. Since the calculations are simple arithmetic, they are readily performed by ENCOA. Assuming that the scores and weights are valid and that the model captures the salient factors relevant to the situation, the course of action yielding the highest weighted score represents the most likely enemy course of action in the opinion of the participating intelligence analysts.

The fifth and final step in utilizing MAUA through computerized decision aids like ENCOA is sensitivity analysis. Different sensitivity analyses can be performed

to systematically test the sensitivity of the total utility scores to variations in the weights assigned to subfactors. Sensitivity analysis can be useful to a group of intelligence analysts, for example, who have conflicting opinions about the importance of particular factors in the hierarchy. The analysis would reveal whether the differences in opinion significantly affect the resultant total utilities of the options. By showing the implications of differences of opinion, MAUA decision aids like ENCOA reduce the emotional aspects of disagreement by promoting a task-focus toward evaluating which differences of opinion truly make a difference overall.

Sage and White (1979) have recently evaluated a MAUA decision aid called RSCREEN for the Command and Control Technical Center, Defense Communications Agency (DCA/C140). RSCREEN is extremely similar in concept to ENCOA. The only difference is in the MAUA hierarchy; all four other steps in RSCREEN are identical to those for ENCOA. <u>The evaluation results strongly supported the general position that decision-analytic techniques and computerized decision aids improve the decision process in general, the judgmental accuracy of individuals in particular, and the communication process of the decision-making group</u>. According to Sage and White (1979, p. 10):

> "RSCREEN was viewed as providing a realistic and adequate situation model structure which encouraged accurate model development for real-world problems. Although the final choice selection by RSCREEN was not always viewed as being consistent with the decision-maker's original choice, there were indications that the aiding process could be convincing enough to change a decision-maker's original choice. . . . Decision process changes were felt by those using the aid to be such as to lead to increased and more effective

thought about the situation of interest, particularly in a group setting. Improved decision-making capability, given sufficient training, experience, and awareness of RSCREEN's situation modeling limitations, were felt to be likely results of decision process changes due to using the aid in appropriate circumstances."

According to Sage and White's evaluation of RSCREEN, the major barrier to the effective utilization of computerized decision aids is the personal decision-making style of key management personnel and institutional constraints. Unless the decision aid fits into both the management style of leadership personnel and the behavioral characteristics of the operational environment, the aid will not be used. Development of computerized decision aids as an adjunct to the proposed IPB Templating process must, therefore, explicitly consider potential user needs and organizational limitations. By so doing, analytically correct aids can be molded to the analysts' environment, thus increasing the probability of their utilization. The interested reader is referred to Adelman, Donnell, Patterson, and Weiss (1980) for a detailed discussion of the importance of user involvement during the development of decision-analytic aids to their successful implementation.

3.2 Research and Aids Applicable to the Revision of Decision Support Templates

The iterative templating process proposed in Circulars TC30-27 and TC34-3 can be represented quantitatively by Bayes' Theorem, which is shown in equation [1].

$$\frac{P(H_1)}{P(H_2)} \quad \times \quad \frac{P(D|H_1)}{P(D|H_2)} \quad = \quad \frac{P(H_1|D)}{P(H_2|D)} \qquad [1]$$

(Prior Probabilities)	(Conditional Probabilities)	(Posterior Probabilities)
Situational Template	Event Templates; Event Analysis Matrices	Decision Templates

The situation template shows how the enemy force would probably look for different courses of action; therefore, it specifies, at a minimum, the initial hypothesis for Bayes' Theorem. If the situation template also provides an initial estimate of the most likely course of action, as we assume it will after analysts have developed the first decision support template, then it specifies the prior probabilities for the hypothesis in some qualitative form.

The event template identifies the information required to support each of the initial hypotheses. Consequently, the information within each event template represents, in qualitative terms, conditional probabilities, for if the enemy actually took a particular course of attack, then certain events have a higher probability of being observed than others. Furthermore, the information within all the event templates represents, again in qualitative terms at present, likelihood ratios [i.e., $P(D|H_1)/P(D|H_2)$] that indicate how much more likely certain events are to be observed than others for one course of attack than another.

Finally, the decision support template represents the intelligence estimate in graphic form. In order to develop this template, the analysts must use collected intelligence data to revise their initial hypotheses about the enemy's most likely course of attack. Consequently, the decision support template represents, again in qualitative terms, the posterior probabilities in Bayes' Theorem.

The next two sections provide a review of research studying Bayesian inference. The goal throughout this review, as in the preceding one, is to identify decision-analytic techniques and computerized decision aids that could be used in conjunction with templating to improve the accuracy of the necessary judgments in the proposed IPB analysis process.

3.2.1 Research findings: without judgment aids - Extensive reviews of psychological research in which subjects' final probability estimates have been compared with those prescribed by Bayes' Theorem can be found in Fischer, Edwards, and Kelly (1978), Rapoport and Wallsten (1972), and Slovic and Lichtenstein (1971). "In general, humans have been found to be very suboptimal processors of probabilistic information. Although they typically revise their opinions in the same direction as Bayes' Theorem, they do not revise them enough" (Fischer et al., p. 6).

This conclusion is based on a variety of experimental tasks in which people were asked to infer which of two or more statistical models (i.e., the hypotheses) generated the data. For example, Phillips and Edwards (1966) used binomial data generators, Phillips, Hays, and Edwards (1966) used multinominal data generators, and DuCharme and Peterson (1968) used normal data generators. This general result of suboptimality in judgment has been called conservatism because people extract less certainty from the data than they should, and consequently, their judgments about the implications of the data are conservative when compared to those of Bayes' Theorem.

This finding could have great implications for IPB, which is essentially a sophisticated Bayesian inference task as we noted previously. For if intelligence analysts using IPB are conservative information processors, then they

are not drawing implications from the data as fast as they could be with Bayes' Theorem. These estimates about enemy courses of action may well be suboptimal because they will not have sufficiently revised their opinions to take full account of the certainty in the data. Consequently, the entire templating process will not convey as much information to commanders as it should. The time available for friendly tactical decision planning and implementation may be reduced considerably if intelligence analysts are conservative information processors.

The potential implications of Bayesian research for IPB are compounded by the fact that intelligence analysts rely on language to convey uncertain information in their event templates and event analysis matrices instead of numerical estimates. It is impossible, however, to directly translate qualitative expressions of uncertainty such as "very likely" into probability values. For while most people would agree that "very likely" means a probability greater than 0.5, there is no general agreement of how much more than 0.5 the probability is (0.8?, 0.9?, 0.99?).

The lack of agreement in the use of language to convey probability estimates has been demonstrated with intelligence analysts in a number of experiments reviewed by Barclay, Brown, Kelly, Peterson, Phillips, and Selvidge (1977) and Phelps, Halpin, Johnson, and Moses (1980). The former, for example, cites an anecdotal study where an intelligence analyst was asked to substitute probability estimates for some of the verbal qualifiers in an article he had written. The first statement was: "The cease-fire is holding but it could be broken within a week." The analyst said that he meant there was a 30% chance the cease-fire would be broken within a week. Later, an analyst who had helped the original analyst prepare the statement said she thought that there was an 80% chance that the cease-fire would be broken.

Yet, both analysts had previously believed that they were in agreement about what could happen.

This anecdote and more importantly, controlled research with analysts, provides strong evidence that significant miscommunication occurs among intelligence analysts. Additional research cited by Phelps et al. (1980, p. 10), "...indicates that a simple change in training procedure, a clarification of scale definitions, etc., would not be adequate to significantly improve the communication of the evaluation..." Seeming agreement among analysts about the implications of data for IPB templates, therefore, might result from the impreciseness of the verbal qualifiers used in the IPB process. In reality, there could be considerable disagreement. When such disagreement finally surfaces, the time available for tactical decision planning and implementation again may be lost as analysts review previously developed templates in an effort to determine the basis of their disagreement.

The fact that verbal qualifiers are imprecise descriptions of levels of certainty provides strong motivation for the use of a quantitative language of uncertainty. Appropriate use of numbers allows uncertainty to be expressed with precision. If an analyst uses numbers, either percentages or odds, to convey a degree of belief about the likelihood of future events, and the numbers are carefully chosen to reflect the analyst's uncertainty, then these numbers can be easily converted into probabilities.

Research is beginning to identify why individuals are conservative information processors. One reason, for example, is that people tend to search for confirming rather than disconfirming evidence of alternative hypotheses (see Einhorn and Hogarth, 1978, for a review). This is often a suboptimal data collection and revision strategy

because data often can confirm many hypotheses at the same time; disconfirming evidence, on the other hand, can quickly eliminate a hypothesis from consideration. Another reason why individuals are conservative information processors is that they tend to use simple rules or "judgment heuristics" (Tversky and Kahneman, 1974) instead of the axioms of probability theory when making subjective probability estimates. As a result of learning why individuals are conservative information processors, researchers have been able to develop decision aids and training methods to help persons make more accurate probability estimates.

A recent review paper by Phelps et al. (1980) suggests that training methods can be developed to improve the accuracy of probability estimates. They cited two experiments by Lichtenstein and Fischhoff (1978) who found that people were able to make more accurate probability judgments after recieving feedback about the direction and magnitude of their initial errors. However, they found that the transfer of this learning to other tasks was moderate, at best. Donnell and DuCharme (1975) have found similar results. As a result, Phelps et al. (1980) suggested that intelligence analysts' training in probability assessment be confined to the tactical intelligence context. This can be readily accomplished, since the training methods used by Lichtenstein and Fischhoff (1978), "...are well documented and could easily be automated for self-instruction and practice" (Phelps et al., 1980, p. 20).

In addition to training methods, analytical decision aids have been developed to help people make more accurate probability estimates. In the next section we discuss the research supporting their development; in a subsequent section we discuss how they might be used to improve IPB.

3.2.2 Research finding: with judgment aids - Efforts to develop judgment aids to help people infer posterior probabilities began in the 1960's with the development of an inference system called PIP--an acronym for Probabilistic Information Processing. In the formulation proposed by Edwards, Lindman, and Phillips (1965), people were tasked with identifying relevant hypotheses, information sources that could discriminate between these hypotheses, and the likelihood ratios linking data with hypotheses. This would be analogous to developing, in more quantitative terms than present, the situation and event templates in IPB. The task of aggregating information across data in PIP was assigned to Bayes' Theorem, since research had indicated that people were conservative information possessors. This would be analogous to using Bayes' Theorem to integrate the intelligence data used to make the intelligence estimate represented graphically in the decision-support template.

Initial efforts by Edwards, Phillips, Hays, and Goodman (1968), Kaplan and Newman (1966), and Wheeler (1972) all found PIP superior to unaided inference; PIP consistently assigned higher posterior probabilities to the true hypotheses. Implicit in the original formulation of PIP was the assumption that the environment could be described in terms of a stationary, single-stage inference model in which all data were conditionally independent of the hypotheses. Subsequent efforts to test PIP varied aspects of this original assumption, thereby representing more complex, yet representative environments.

Research in more complex environments also supported the value of PIP, but the results were less clear. For example, Domas and Peterson (1972) found PIP to outperform unaided judgment when data were conditionally independent, but not when the data were conditionally dependent,

for in the latter case PIP assigned excessively high probabilities. In contrast, Schum, Southard, and Wombolt (1969) found a modified PIP, termed semi-PIP, to perform quite well with conditionally dependent data. In semi-PIP, persons first grouped data into bundles that were conditionally independent, although the data within each bundle was conditionally dependent. Persons then assessed the likelihood ratios for the conditionally independent bundles, and Bayes Theorem combined these estimates into posterior probabilities. Semi-PIP was found to be substantially better than unaided judgments, particularly as the number of data to be aggregated increased.

PIP Systems have also been developed to deal with the problem of hierarchical (also called multistage or cascaded) inference. Hierarchical inference problems involve several levels of analysis and therefore, are representative of many intelligence problems. Experimental studies of intuitive hierarchical inference by Gettys, Kelly, and Peterson (1973) and Schum, DuCharme, and DePitts (1971) have also found humans to be suboptimal information processors. The reason for this, however, is that their posterior probabilities were too extreme when compared to those of Bayes' Theorem. To the extent that IPB incorporates hierarchical inference, it is susceptible to yet another limitation in unaided judgment.

Kelly and Barclay (1973) provide general mathematical models for hierarchical inference that can be directly translated into computer algorithms for decision aids. And Barclay (1976) describes an interactive graphics aid for Bayesian hierarchical inference. From a practical standpoint, however, these models are tractable only when the problem can be structured to eliminate most conditional

dependencies between data. If this cannot be done, individuals would be required to make a prohibitively large number of judgments.

Experimental evidence on hierarchical PIP systems is unfortunately scant. Gettys, Kelly, Peterson, Michel, and Steiger (1973) have conducted two relevant studies, however, both of which demonstrated the superiority of a hierarchical PIP system over unaided inference.

Hierarchical PIP systems (called Bayesian hierarchical inference models) have been developed for a number of strategic intelligence problems. For example, Barclay, Kelly, and Stewart (1976) developed a hierarchical PIP system for assessing the personnel strength of foreign ground force units. And Peterson, Randall, Shawcross, and Ulvila (1976) developed a hierarchical PIP system for the Navy to predict an anti-ship missile threat. Since the PIP system developed by Barclay et al. (1976) is classified, we will discuss the PIP system developed by Peterson et al. (1976) to better help the reader understand its conceptual framework.

The structure of the Bayesian hierarchical model developed for a Red air/submarine threat is presented in Figure 3-4. In this model, a series of different enemy activities, and the associated observations that might be available to a friendly task force commander, were postulated. Then conditional probabilities were assessed linking the activities to the hypotheses of routine surveillance, feint attack, and attack, and linking the observations to the activities. For example, consider the targeting activity. It was hypothesized that Red's targeting activity might occur in one of three ways: intermittently, steadily, or not at all. If Red intended no attack, thus supporting the "routine" hypothesis, it was most likely that Red would

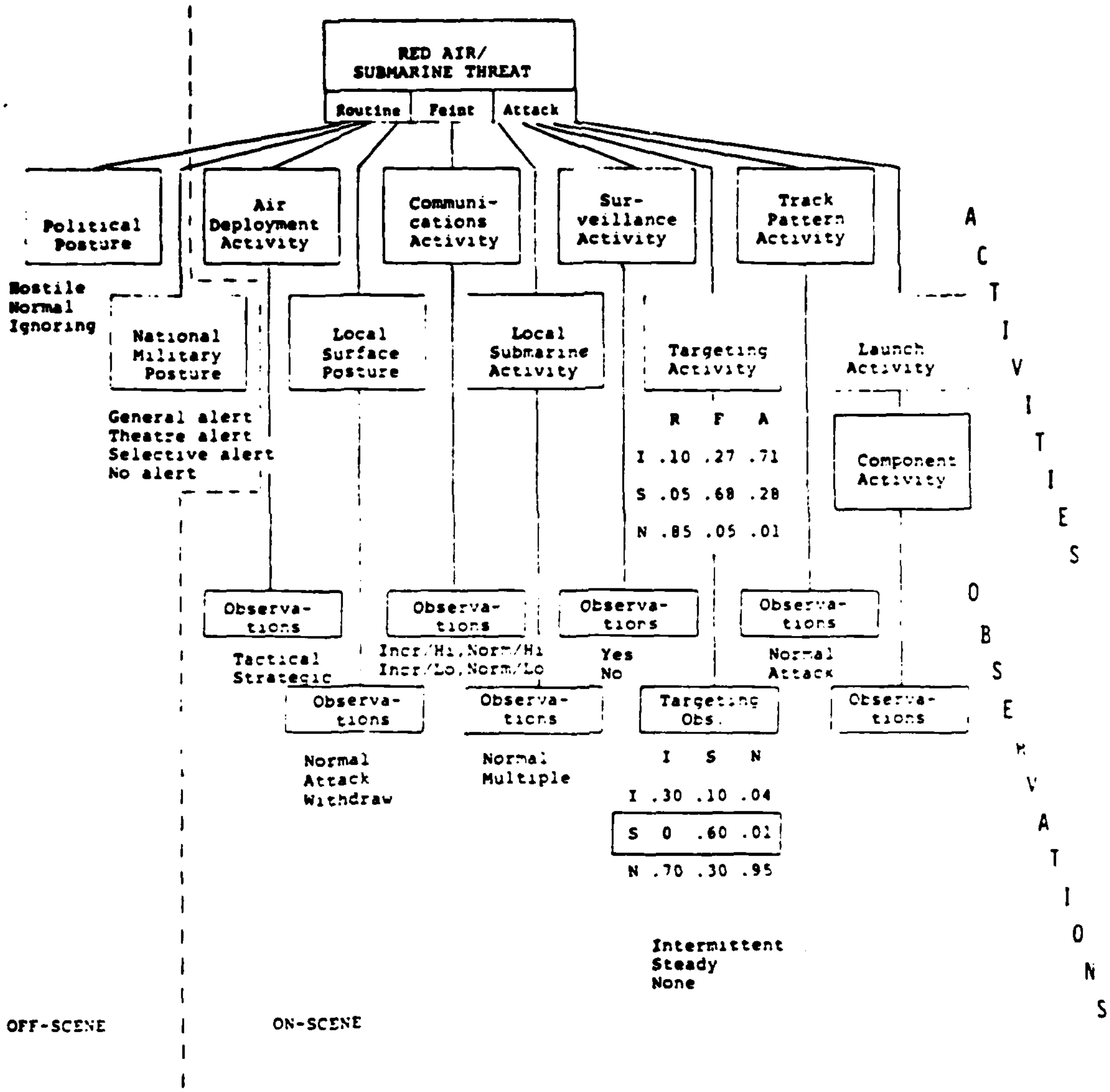

Figure 3-4

BAYESIAN HIERARCHICAL INFERENCE MODEL FOR STRATEGIC INTELLIGENCE EXAMPLE

not engage in any targeting. Consequently, conditional probabilities of .10, .05, and .85 were assessed for the activities of intermittent, steady, and none, respectively, given the "routine" hypothesis, as shown in the top matrix in Figure 3-4. If Red intended feint, it was thought that, most likely, Red would target in a steady fashion. Consequently, conditional probabilities of .27, .68, and .05 were assessed for intermittent, steady, and none, respectively. Finally, if Red intended to attack, it was most likely that Red would try to conceal the target signal by using it intermittently; consequently, conditional probabilities of .71, .28, and .01 were assessed for intermittent, steady, and none.

In a similar manner, the bottom matrix of Figure 3-4, linking observations to activities, was assessed. If Red actually was targeting intermittently (an activity), it was most likely that the task force commander would receive a report that Red was not targeting (an observation). Thus, conditional probabilities of .30, 0, and .70 were assessed for the observation of intermittent, steady, and none, respectively. Further, conditional probabilities of .10, .60, and .30 were assessed for the observation of intermittent, steady, and none given that Red's activity was steady targeting and probabilities of .04, .01, and .95 for observations of intermittent, steady, and none if Red was not targeting.

All the conditional probabilities in the hierarchical model shown in Figure 3-4 had to be assessed by people. Once the conditional probabilities have been entered, however, they are combined analytically (arithmetically), not intuitively, in the PIP system. Specifically, by multiplying the lower matrix of Figure 3-4, which links the observations to the activities, by the upper matrix, which links the activities to the hypotheses, yields a

product matrix that relates the observations to the hypotheses. Table 3-1 shows this multiplication for the targeting example explained above. This product matrix provides output in the desired form of an indicator list relating observations to hypotheses, in terms of the probability that the former implies the latter. As a result, the overall probability of each hypothesis (i.e., possible enemy intent) now can be determined explicitly on the basis of the observations made at a particular time.

As this example illustrates, hierarchical PIP systems have been developed to help people make complex probability judgments. We now turn to consider directly the applicability of general PIP decision aids in IPB.

3.2.3 A judgment aid - It is important to reiterate at the outset that the judgment process supporting the revision of templates within IPB can be conceptualized in terms of the elements of Bayes' Theorem. The decision support template at the end of one iteration is the situation template at the beginning of the next iteration. The situation template is represented by the prior probabilities, which indicate the relative likelihood of the different OPFOR courses of action (i.e., hypotheses) under consideration. The event templates and decision matrices are represented by the conditional probabilities, which indicate the relative likelihood that certain events support particular courses of action. The posterior probabilities are represented by the decision support template, which indicates the revised likelihood of the courses of action (i.e., hypotheses) on the basis of observed data. This new estimate of enemy intent is then input to friendly tactical decision making and subsequent action.

The literature cited previously indicates that people have considerable difficulty in making judgments

$$
\begin{array}{cc}
 & \text{Targeting Activities} \\
 & \begin{array}{ccc} \text{Inter.} & \text{Steady} & \text{None} \end{array} \\
\text{Targeting Observations} \;
\begin{array}{l} \text{Intermittent} \\ \text{Steady} \\ \text{None} \end{array} &
\left[\begin{array}{ccc} .30 & .10 & .04 \\ 0 & .60 & .01 \\ .70 & .10 & .95 \end{array}\right]
\end{array}
\times
\begin{array}{cc}
 & \text{Threat Hypotheses} \\
 & \begin{array}{ccc} \text{Routine} & \text{Feint} & \text{Attack} \end{array} \\
\text{Targeting Activities} \;
\begin{array}{l} \text{Intermittent} \\ \text{Steady} \\ \text{None} \end{array} &
\left[\begin{array}{ccc} .10 & .27 & .71 \\ .05 & .68 & .28 \\ .85 & .05 & .01 \end{array}\right]
\end{array}
=
\begin{array}{cc}
 & \text{Threat Hypotheses} \\
 & \begin{array}{ccc} \text{Routine} & \text{Feint} & \text{Attack} \end{array} \\
\text{Targeting Observations} \;
\begin{array}{l} \text{Intermittent} \\ \text{Steady} \\ \text{None} \end{array} &
\left[\begin{array}{ccc} .07 & .15 & .24 \\ .04 & .41 & .17 \\ .89 & .44 & .59 \end{array}\right]
\end{array}
$$

Table 3-1

MATRIX MULTIPLICATION RELATING OBSERVATION
TO ENEMY INTENT (STRATEGIC INTELLIGENCE EXAMPLE)

consistent with Bayes' Theorem. People extract less certainty from the data than they should, and consequently, their judgments about the implications of the data are conservative when compared to those of Bayes' Theorem. Since the judgment process supporting the revision of IPB templates can be conceptualized as a Bayesian inference process, this finding could have great implications for IPB. For if intelligence analysts using IPB are conservative information processors, then they are not drawing implications from the data as fast as they could be with Bayes' Theorem. Their estimates about enemy courses of action may well be suboptimal because they will not have sufficiently revised their opinions to take full account of the certainty in the data. Consequently, the entire templating process will not convey as much information to commanders as it should. As a result, force commanders will lose time for tactical decision making.

General decision aids, called Probabilistic Information Processing (PIP) systems have been developed to ensure that human judgment is consistent with Bayes' Theorem for a number of intelligence problems (e.g., see Peterson et al., 1976). Such a decision aid is now being developed to help tactical intelligence analysts revise their judgment about the most likely OPFOR course of action on the basis of new information. This aid is being designed so that it can be readily integrated into the proposed iterative IPB analysis process. Consequently, it will provide an important adjunct to the development and utilization of IPB templates. The decision aid is scheduled for initial evaluation during 1981.

The Bayesian judgment aid will probably operate in the following manner:

(1) the analysts define the n different enemy courses of action (COAs) under consideration (i.e., enter a brief [< 10 character title]);

(2) the analysts enter a set of prior probabilities for the n potential courses of action;

(3) for a given datum, the analysts input a brief title and the probability of that datum conditional upon each COA being considered;

(4) the analysts inspect the posterior probabilities (or likelihood ratios);

(5) they revise any posterior probabilities that are counter-intuitive; and

(6) the analysts report on potential enemy COAs based on the probabilities after Step 5 or return to Step 3 if there are additional data.

Step 5 is necessary since any redundancy or facilitation in the data is not likely to be taken into account in providing the probability estimates in Step 3. For the probabilities in Step 4 to always be correct, without some revision at Step 5, all data must be independent. An alternative to Step 5 would be to require that the probabilities estimated in Step 3 be conditional not only on the hypothesized COA but also upon all data that have come before. This type of probability estimate is quite difficult, however, and is most easily performed using a Bayesian hierarchical (staged) inference (BHI) model rather than the simple model proposed for use here.

The ability to use a Bayesian framework to represent the different judgments inherent in revising the IPB templates is illustrated in the above steps. Step 1 represents the different OPFOR courses of action represented in the situation template. Step 2 represents the relative likelihood of these actions at the end of one iteration of the IPB analysis process. Step 3 represents the judgments in the event templates and event matrices, which indicate the relative likelihood of having collected the newly acquired data on the basis of the enemy actually taking different courses of action. Steps 4, 5, and 6 represent the intelligence analysts' estimates of enemy intent, as represented graphically in the decision-support template.

It is important to note that a Bayesian aid also can be used even if there is no revision of IPB templates. In this case, the prior probabilities are assumed to be equal for each enemy course of action unless the analysts think certain actions are more likely than others (Step 2). The analysts then either estimate (1) the probability of the individual intelligence datum conditional on each of the hypothesized courses of action as the datum is collected (Step 3), or (2) the probability of all collected intelligence data conditional on each action. In the former case, Bayes' Theorem generates posterior probabilities after collecting each datum; the posterior probabilities are then the prior probabilities that are revised on the basis of the next datum. In the latter case, Bayes' Theorem generates the final posterior probabilities. The former case is recommended in order to minimize conservatism in the analysts' final estimate of enemy intent. Notice that while the posterior probabilities are constantly revised on the basis of new information, the decision-support template is not because it is not developed until the analysts have collected the data necessary for the final estimate of enemy intent.

ENCOA also can be used for revising decision support templates on the basis of collected intelligence data. During the next iteration, the analysts modify (1) the score of each OPFOR course of action on each of the attributes in the hierarchy to represent the implications of collected data, and (2) the weights on the attributes at each level, if the collected data has altered the relative importance of going from the lowest to the highest scores on the attributes. The scores and weights are then combined to indicate the most likely enemy course of action. Again, sensitivity analysis provides a means for determining the effect of using different scores and weights on overall judgments of enemy intent. The overall utility scores for each OPFOR course of action can be readily represented in graphical form in the next decision-support template.

An advantage of using Bayes' Theorem instead of multi-attribute utility theory to represent quantitatively the judgments inherent in template revision is that the posterior probabilities can be readily incorporated into an expected utility theoretic framework for friendly tactical decision making. This framework is represented pictorially in the decision tree shown in Figure 3-5. In this representation, the friendly commander can take either action #1 or action #2. Each action will have a particular outcome (or consequence), depending on the course of action actually adopted by the enemy. Furthermore, each outcome will vary in utility (or value) to the friendly commander. Actual enemy intent is, of course, assumed to be unknown at the time of the decision.

Bayes' Theorem can be used to indicate the probability (or likelihood) that the enemy is taking course of action A or B on the basis of collected data. This information can be represented pictorially in the decision-

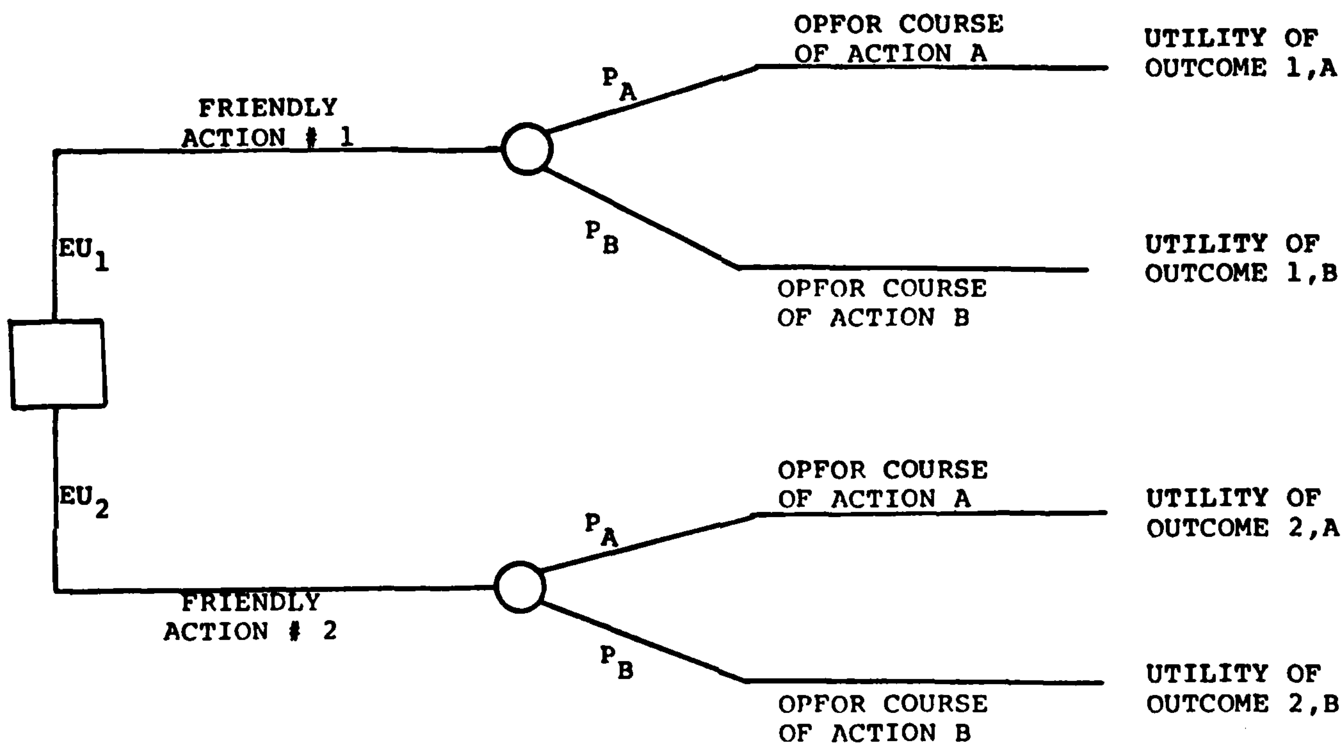

Figure 3-5

DECISION TREE EXAMPLE REPRESENTING THE APPLICABILITY OF EXPECTED UTILITY FRAMEWORK FOR TACTICAL DECISION MAKING

support template presented to the friendly commander. In addition, Bayes' Theorem can be used to analytically revise the relative likelihood estimates for different enemy courses of action (e.g., A and B in Figure 3-3) on the basis of newly collected data organized in the decision matrices proposed in Circulars TC30-27 and TC34-3. Again, the relative likelihood estimates can be presented to the friendly commander in the decision-support template.

Expected utility theory provides an analytic procedure for combining these relative likelihood estimates with the friendly commanders' estimates of the utility of different outcomes resulting from their and the enemy's actions. In terms of Figure 3-3, the friendly commander should, first, multiply the utility and probability estimate for each of the four branches of the decision tree shown in Figure 3-3, and then, sum the two resulting values for each action to determine the expected utility of each action. The commander then should select the action with the highest expected utility in order to ensure the greatest degree of success over time. In this fashion, the posterior probability judgments of enemy intent that are inherent in template revision can be readily incorporated into friendly tactical decision making through an explicit, retraceable analysis process.

3.3 Summary

IPB depends on the subjective judgment of intelligence analysts. These judgments can be represented conceptually in terms of a causal inference model, for analysts must infer enemy intent on the basis of tactical indicators and information. Research on causal inference with experts in different fields indicates that the level of judgmental accuracy and interpersonal agreement is a direct function

of the characteristics of the judgmental task facing the experts. This conclusion is further supported by controlled laboratory research, which has shown (1) that judgmental accuracy and agreement can be increased or decreased by manipulating characteristics of the task, and (2) that relatively high levels of accuracy and agreement can be maintained under conditions that normally prevent it, by using judgment aids that provide persons with information about task characteristics.

This section describes two judgment aids for improving the judgmental process inherent in IPB. The first aid uses multi-attribute utility assessment techniques to facilitate the development of the initial decision-support template. The judgment aid is called ENCOA, for Enemy Courses of Action. ENCOA provides analysts with a systematic procedure for evaluating each potential OPFOR course of action on twenty-four factors affecting enemy intent. In addition, ENCOA can be used to present quantitatively and pictorially the scores, weights, overall utilities, and subsequent sensitivity analyses to the friendly commander. Such information describing (1) how the analysts reached their conclusions about the most likely OPFOR course of action, as well as (2) the implications of differences in opinion between the analysts, represent an important adjunct to the graphic decision-support template.

The second decision aid uses Bayesian assessment techniques to facilitate the revision of decision-support templates on the basis of newly collected data represented in the event templates and event analysis matrices. The iterative process for revising IPB templates can be represented quantitatively by Bayes' Theorem. Psychological research comparing subjects' final probability estimates with those prescribed by Bayes' Theorem has shown that subjects do not

revise their final probability estimates far enough; they are conservative. Therefore, Bayesian judgment aids, called Probabilistic Information Processing systems, should help analysts draw better implications from collected intelligence data. These implications, represented as probability estimates, can then be incorporated into the event and decision support templates. Consequently, such aids should help analysts implement the template revision process outlined in Circulars TC30-27 and TC34-3.

4.0 CONCLUSION

Intelligence Preparation of the Battlefield represents a major step toward improving procedures for correlating data with enemy intent. However, it can be improved, for the process whereby input (new intelligence data) is translated into output (templates) necessarily depends on analysts' judgments about the implications of multiple pieces of potentially fallible data. Except for terrain analysis (Step 3) and weather analysis (Step 4), Circulars TC30-27 and TC34-3 fail to tell intelligence analysts how they are to make the judgments necessary to implement the proposed IPB analysis process. Furthermore, they provide little information on how analysts are to revise event and decision-support templates on the basis of new information. Finally, they fail to describe decision-analytic techniques and judgmental aids that are available (or being developed) to help analysts make the necessary judgments. Scientific research on causal inference strongly suggests that such techniques and aids would improve (1) the cognitive skills and, thus, the judgmental accuracy of intelligence analysts developing IPB templates, and (2) the communication process for conveying the reasons for these judgments to tactical commanders.

REFERENCES

Adelman, L.; Donnell, M. L.; Patterson, J. F.; and Weiss, J. J. Issues in the Design and Evaluation of Decision-Analytic Aids. Technical Report TR 81-1-304. McLean, VA: Decisions and Designs, Inc., February 1981.

Adelman, L., and Mumpower, J. L. The analysis of expert judgment. Technological Forecasting and Social Change, 1979, 15, 191-204.

Adelman, L.; Donnell, M. L.; and Phelps, R. H. Judgment-analytic Intelligence Preparation of the Battlefield: Critique and Recommendations. Technical Report PR 81-5-304. McLean, Virginia: Decisions and Designs, Inc., 1980, in press.

Barclay, S. Interactive Graphical Aids for Bayesian Hierarchical Inference. Technical Report 76-9. McLean, Virginia: Decisions and Designs, Inc., August 1976. AD A034598

Barclay, S.; Brown, R. V.; Kelly, C. W. III; Peterson, C. R.; Phillips, L. D.; and Selvidge, J. Handbook for Decision Analysis. Technical Report TR 77-6-30. McLean, Virginia: Decisions and Designs, Inc., September 1977. AD A049221

Barclay, S.; Kelly, C. W., III; and Stewart, R. R. A Formal Process for Assessing the Personnel Strength of Foreign Ground force Units (U). (XGDS) Final Technical Report (S/NF). McLean, Virginia: Decisions and Designs, Inc., January 1976.

Bowen, R. J.; Halpin, J. A.; Russell, P. T.; and Staniforth, B. J. Tactical order of battle: A State-of-the-art survey. Technical Paper No. 265. Alexandria, Virginia: U.S. Army Research Institute for the Behavioral and Social Sciences, October, 1975.

Brehmer, B. Policy conflict as a function of policy differences and policy complexity. Scandinavian Journal of Psychology, 1972, 13, 208-221.

Brehmer, B. Effect of cue intercorrelations on interpersonal learning of probabilistic inference tasks. Organizational Behavior and Human Performance. 1974, 12, 397-412.

Brehmer, B. Social judgment theory and the analysis of interpersonal conflict. Psychological Bulletin, 1976, 83, 985-1003.

Brunswik, E. The conceptual framework of psychology. International Encyclopedia of Unified Science. Vol. I, No. 10, Chicago: University of Chicago Press, 1952.

Brunswik, E. Perception and the representative design of psychological experiments. (2nd Ed.) Berkely and Los Angeles: University of California Press, 1956.

Cook, R. L., and Stewart, T. R. A comparison of seven methods for obtaining subjective descriptions of judgmental policy. Organizational Behavior and Human Performance, 1975, 13, 31-45.

Domas, P. A., and Peterson, C. R. Probabilistic information processing systems: Evaluation with conditionally dependent data. Organizational Behavior and Human Performance, 1972, 7, 77-85.

Donnell, M. L., and DuCharme, W. M. The effect of Bayesian feedback on learning in an odds estimation task. Organizational Behavior and Human Performance, 1975, 14, 305-313.

DuCharme, W. M., and Peterson, C. R. Intuitive inference about normally distributed populations. Journal of Experimental Psychology, 1968, 78, 269-275.

Dudycha, A. L., and Naylor, J. C. Characteristics of the human inference process in complex choice behavior situations. Organizational Behavior and Human Performance, 1966, 1, 110-128.

Ebbesen, E., and Konecni, V. Decision making and information integration in the courts. Journal of Social and Personality Psychology, 1975, 32, 805-821.

Ebert, R. J., and Kruse, T. E. Bootstrapping the security analyst. Journal of Applied Psychology, 1978, 63, 110-119.

Edwards, T.; Lindman, H.; and Phillips, L. D. Emerging technologies for making decisions. In New Directions in Psychology II. Holt, Rinehart, and Winston, 1965, 261-325.

Edwards, W.; Phillips, L. D.; Hays, W. L.; and Goodman, B. C. Probabilistic information processing systems: Design and evaluation. IEEE Trans. Syst. Sci., Cybernetics, 1968, 248-265.

Einhorn, H. J. Expert Judgment: Some necessary conditions and an example. Journal of Applied Psychology, 1974, 59, 562-571.

Einhorn, H. J., and Hogarth, R. M. Confidence in judgment: Persistence of the illusion of validity. Psychological Review, 1978, 85, 395-416.

Fischer, G. W.; Edwards, W.; and Kelly, C. W., III. Decision Theoretic Aids for Inference, Evaluation, and Decision Making: A Review of Research and Experience. Technical Report TR 78-1-30. McLean, Virginia: Decisions and Designs, Inc., February 1978. AD A053962.

Gettys, C. F.; Kelly, C. W., III; and Peterson, C. R. The best guess hypothesis in multi-stage inference. Organizational Behavior and Human Performance, 1973, 10, 364-373.

Gettys, C. F.; Kelly, C. W, III; Peterson, C. R.; Michel, C.; and Steiger, J. H. Multiple-stage probabilistic information processing. Organizational Behavior and Human Performance, 1973, 10, 374-387.

Goldberg, L. R. Man versus model of man: A rationale, plus some evidence, for a method of improving on clinical inferences. Psychological Bulletin, 1970, 73, 422-432.

Hammond, K. R. Computer graphics as an aid to learning. Science, 1971, 172, 903-908.

Hammond, K. R., and Adelman, L. Science, values and human judgment. Science, 1976, 194, 389-396.

Hammond, K. R., and Boyle, P. J. Quasirationality, quarrels, and new conceptions of feedback. Bulletin of the British Psychological Society, 1971, 24, 103-113.

Hammond, K. R., and Brehmer, B. Quasirationality and distrust: Implications for international conflict. In L. Rappoport and D. A. Summers (Eds.), Human Judgment and Social Interaction. New York: Holt, Rinehart, and Winston, 1973.

Hammond, K. R.; Hursch, C. J.; and Todd, F. J. Analyzing the components of clinical inference. Psychological Review, 1964, 71, 438-456.

Hammond, K. R.; Rohrbaugh, J.; Mumpower, J.; and Adelman, L. Social judgment theory: Applications in policy formation. In M. F. Kaplan and S. Schwartz (Eds.), Human judgment and decision processes: Applications in problem settings. New York: Academic Press, 1977.

Hammond, K. R.; Stewart, T. R.; Brehmer, B.; and Steinmann, D. O. Social judgment theory. In M. F. Kaplan and S. Schwartz (Eds.), Human judgment and decision processes: Applications in problem settings. New York: Academic Press, 1975.

Hoffman, P. J.; Slovic, P.; and Rorer, L. G. An analysis of variance model for the assessment of configural cue utilization in clinical judgment. Psychological Bulletin, 1968, 69, 338-349.

Hursch, C. J.; Hammond, K. R.; and Hursch, J. L. Some methodological considerations in multiple-cue probability studies. Psychological Review, 1964, 71, 42-60.

Johnson, E. M.; Spooner, R. L.; and Jaarsma, D. The perception of tactical intelligence indications by intelligence officers. Technical Paper 278. Alexandria, Virginia: U.S. Army Research Institute for the Behavioral and Social Sciences, April, 1977.

Kaplan, R. J., and Newman, J. R. Studies in probabilistic information processing. IEEE Trans. Hum. Fact. Electronics, 1966, 7, 49-63.

Kelly, C. W. Program Completion Report: Advanced Decision Technology Program (1972-1979). Technical Report TR 79-3-93. McLean, Virginia: Decisions and Designs, Inc., 1979.

Kelly, C. W., III, and Barclay, S. A general bayesian model for hierarchical inference. Organizational Behavior and Human Performance, 1973, 19, 388-403.

Kibler, A. W.; Watson, S. R.; Kelly, C. W., III; and Phelps, R. H. A prototype aid for evaluating alternative courses of action for tactical engagement. Technical Paper TR-78-A38. Alexandria, Virginia: U.S. Army Research Institute for the Behavioral and Social Sciences, November, 1978.

Lichtenstein, S., and Fischhoff, B. Training for calibration. Technical Report TR-78-A32. Alexandria, Virginia: U.S. Army Research Institute for the Behavioral and Social Sciences, November, 1978.

Lindell, M. K. Cognitive and outcome feedback in multiple-cue probability learning tasks. Journal of Experimental Psychology, 1976, 2 (6), 739-745.

Meehl, P. E. Clinical versus statistical prediction: A theoretical analysis and a review of the evidence. Minneapolis: University of Minnesota Press, 1954.

Mumpower, J. L., and Hammond, K. R. Entangled task-dimensions: An impediment to interpersonal learning. Organizational Behavior and Human Performance, 1974, 11, 377-389.

Patterson, J. F., and Phelps, R. H. Computerized procedures guide: Decision aid for evaluating enemy courses of action (ENCOA). Working Paper Draft. Alexandria, Virginia: U.S. Army Research Institute for the Behavioral and Social Sciences, January, 1980.

Peterson, C. R.; Randall, L. S.; Shawcross, W. H.; and Ulvila, J. W. Decision Analysis as an Element in an Operational Decisions aiding System (Phase III). Technical Report 76-11. McLean, Virginia: Decisions and Designs, Inc., October 1976. AD A034037.

Phelps, R. H.; Halpin, S. M.; Johnson, E. M.; and Moses, F. L. The use of subjective probability in Army intelligence estimation and data evaluation. Working Paper Draft. Alexandria, Virginia: U.S. Army Research Institute for the Behavioral and Social Sciences, May, 1980.

Phelps, R. H., and Shanteau, J. Livestock judges: How much information can they use? Organizational Behavior and Human Performance, 1978, 21, 209-219.

Phillips, L. D., and Edwards, W. Conservatism in a simple probability inference task. Journal of Experimental Psychology, 1966, 72, 346-357.

Phillips, L. D.; Hays, W. L.; and Edwards, W. Conservatism in complex probabilistic inference. IEEE Trans. Hum. Fact. Electronics, 1966, 7, 7-18.

Rapoport, A., and Wallsten, T. S. Individual decision behavior. Annual Review of Psychology, 1972, 23, 131-176.

Sage, A. P., and White, C. C., III. Evaluation of Two DDI Decision Aids Developed for DCA:C140. Document Number 33737-W114-RU-00. Falls Church, Virginia: TRW Defense and Space Systems Group, January 1980.

Schmitt, N.; Coyle, B. W.; and King, L. Feedback and task predictability as determinants of performance in multiple cue probability learning tasks. Organizational Behavior and Human Performance, 1976, 16, 388-402.

Schum, D. A.; DuCharme, W. M.; DePitts, K. E. Research on human multi-stage inference processes. Rice University, Applied Mathematics and Systems Theory, No. 46-11, 1971.

Schum, D. A.; Southard, J.; and Wombolt, L. Aided human processing of inconclusive evidence in diagnostic systems: a summary of experimental evaluations. AMRL-TR-69-11, Aerospace Medical Research Laboratory, Wright-Patterson AFB, 1969.

Slovic, P.; Fischhoff, B.; and Lichtenstein, S. Behavioral decision theory. Annual Review of Psychology, 1977, 28, 1-39.

Slovic, P.; Fleissner, D.; and Bauman, W. S. Analyzing the use of information in investment decision making: A methodological proposal. Journal of Business, 1972, 35, 283-301.

Slovic, P., and Lichtenstein, S. Comparison of Bayesian and regression approaches to the study of information processing in judgment. Organizational Behavior and Human Performance, 1971, 6, 649-744.

Summers, D. A.; Taliaferro, J. D.; and Fletcher, D. J. Subjective versus objective descriptions of judgment policy. Psychonomic Science, 1970 a, 18, 249-250.

Summers, D. A.; Taliaferro, J. D.; and Fletcher, D. J. Judgment policy and interpersonal learning. Behavioral Science, 1970 b, 15, 514-521.

Todd, F. J., and Hammond, K. R. Differential feedback in two multiple-cue probability learning tasks. Behavioral Science, 1965, 10, 429-435.

Tucker, L. R. A suggested alternative formulation in the development by Hursch, Hammond, and Hursch, and Hammond, Hursch, and Todd. Psychological Review, 1964, 71, 528-530.

Tversky, A., and Kahneman, D. Judgment under uncertainty. Heuristics and biases. Science, 1974, 185, 1124-1131.

Uhl, C. N. Learning of interval concepts: 1. Effects of differences in stimulus weights. Journal of Experimental Psychology, 1963, 66, 264-273.

Uhl, C. N. Effects of multiple stimulus validity and criterion dispersion on learning of interval concepts. Journal of Experimental Psychology, 1966, 72, 519-527.

Wheeler, G. Misaggregation versus response bias as explanations for conservative inference. Doctoral Dissertation, University of Michigan, 1972.